Intel® Cloud Builders Guide: Cloud Design and Deployment on Intel® Platforms

NetApp* Unified Networking and Storage: 10 GbE FCoE and iSCSI

Intel® Xeon® Processor 5500 Series
Intel® Xeon® Processor 5600 Series

AUDIENCE AND PURPOSE

This paper is for IT organizations and service providers who are interested in building cloud data centers. Storage solutions are critical elements of the overall cloud architecture. Unified networking based on 10 Gigabit Ethernet provides a cost-effective solution for cloud storage architecture based on commonly used Ethernet technology.

The purpose of this paper is to give complete architectural details with step-by-step instructions to set up and evaluate the two most common block-level storage protocols, fibre channel over Ethernet (FCoE) and Internet small computer system interface (iSCSI), run over 10 GbE from end-to-end. We provide our observations about the functionality and performance observed during our tests.

Since there are tradeoffs among the protocol choices, we recommend that the reader take the time to evaluate the multiple options that exist for unified networking in their own labs running their own workloads to best determine the most appropriate technology choice.

Table Of Contents

Executive Summary

In recent years, growth in Ethernet-based storage has surpassed that of storage-specific fabrics, driven in large part by the increase in server virtualization. Unified networking over 10 Gigabit Ethernet (10 GbE) in the data center offers compelling benefits for the evolving virtualized data center, including simplified infrastructure, lower equipment and power costs, and greater flexibility. Easy access to the storage area network (SAN) will be critical as virtualization deployments continue to grow and new, on-demand data center models emerge.

Intel and NetApp are two leaders in the transition to 10 GbE unified networking. This paper explores the benefits of unified networking, approaches to enable it, and methods Intel and NetApp use to bring important, consolidation-driven technologies to enterprise data center customers.

Introduction

Our Test Plan and Results

Intel and NetApp worked together to implement both FCoE and iSCSI over an end-to-end 10 GbE network in an Intel® lab. In all cases, we used native OS initiators with a standard Intel® 10 Gb Ethernet NIC on the server, rather than offloading the entire protocol stack to a proprietary host bus adapter (HBA) or converged network adapter (CNA) as was done in the past. We were able to achieve near 10 Gbps line rate throughput performance in most cases. We obtained similar results for FCoE and iSCSI in most cases. We present the detailed technology overview, test plan, step-by-step configuration, results, and analysis below. We hope our work will bring new insights to your organization on the use of unified networking over 10 GbE and the use of native OS initiators for iSCSI & FCoE.

The 10 Gigabit Ethernet Transition

Simplification of the Network with 10 GbE

As IT departments look to reduce costs and improve server efficiency, they turn increasingly to server virtualization and consolidation. The benefits of virtualization are widely understood: less server hardware to purchase, lower power and cooling needs, and centralized management. Today's servers are based on powerful new processors, including the Intel® Xeon® processor 5600 and 7500 series, that support more virtual machines (VMs) per physical host than ever before, which helps IT realize greater consolidation ratios.

- The latest generation of Intel® Xeon® processors enables IT to consolidate servers at a 15:1 ratio, which delivers power savings of up to 90 percent and a five-month return on investment.
- New four-socket processors deliver 20 times the performance of previous generation processors.
- Nearly 50 percent of the four-socket servers shipped today are used for virtualization.[1]

Unfortunately, the success achieved by many organizations as they attain these benefits has been limited by past practices used to network virtualized servers. As VM density increases, a physical server's networking needs also increase, which adds both cost and complexity. A typical virtualized server contains eight to ten Gigabit Ethernet (GbE) LAN ports and two dedicated SAN ports.

As server virtualization continues to spread, 10 GbE unified networking simplifies server connectivity. The consolidation of multiple GbE connections traffic onto a single 10 GbE adapter significantly reduces cable and infrastructure complexity, and overall TCO. Enhancements to the Ethernet standard also enable 10 GbE support for both LAN and SAN traffic, which allows IT to realize further benefits as it converges data and storage infrastructures. Thanks to Ethernet's ubiquity, cost-effectiveness, flexibility, and ease of use, it has emerged as the data center unified fabric.

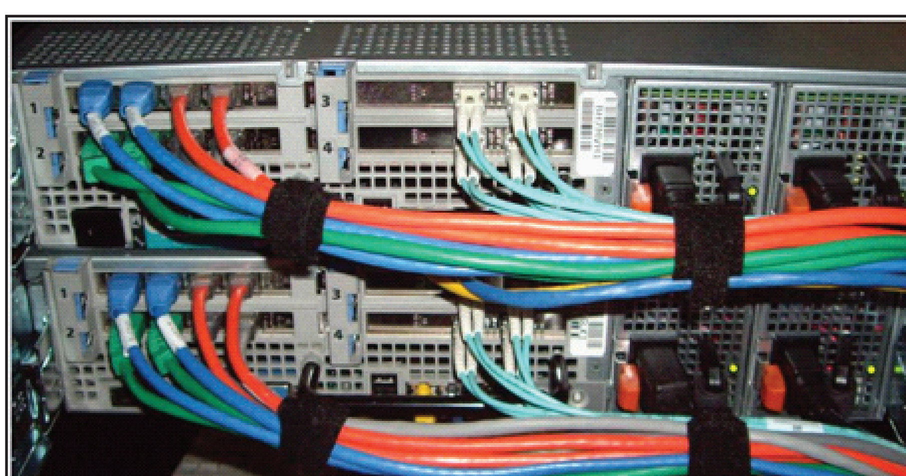


Figure 1. Multiple GbE Connections in a Typical Virtualized Server

10 GbE: The Fabric of Cloud Computing

The growth in server virtualization has helped data center networks evolve from discrete, siloed infrastructures to more flexible fabrics with the scalability and agility necessary to address the needs of new usage models and to provide an excellent foundation for enterprise cloud computing.

Over 2.5 billion users will connect to the Internet in the next five years,[2] with over 10 billion devices.[3] This usage will require eight times the current amount of storage capacity, 16 times the current network capacity, and over 20 times the current compute capacity.[4] A new infrastructure must emerge to power this growth and to enable the most efficient use of resources; this is cloud computing. The cloud is an evolution of computing that delivers services over the Internet to consumers and enterprises. Services scale—as needed and only when needed—without user intervention. We need highly scalable and efficient cloud architecture to provide both the technical attributes, and the extreme resource utilization and efficiency cloud computing promises.

With its reduced hardware requirements, fewer points of management, and broad ecosystem support, 10 GbE delivers the flexible, simplified network infrastructure needed to support cloud computing. The following key characteristics make 10 GbE the ideal fabric for cloud infrastructures.

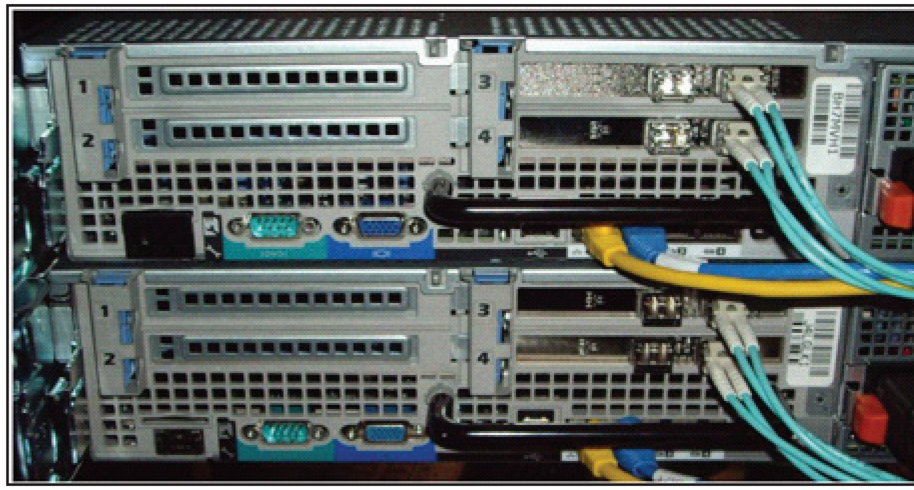


Figure 2. Simplified Server Connectivity with 10 GbE

Ubiquity: Ethernet connectivity ships standard on nearly every server today, and Ethernet infrastructures are a universal data center component. When 10 GbE LAN on motherboard (LOM) connections are integrated in the next generation of servers, unified LAN and SAN connectivity will be available by default.

Advanced Virtualization Support: Advanced server virtualization enables dynamic resource allocation and every cloud computing infrastructure requires it. Technologies from companies such as Intel, VMware, and Microsoft deliver line-rate 10 GbE throughput and support for platform virtualization enhancements.

Unified Networking: A 10 GbE unified fabric simplifies the network infrastructure because it consolidates LAN and SAN traffic. Recent Ethernet enhancements ensure quality of service (QoS) for critical traffic.

Intel and NetApp are companies leading the shift to 10 GbE-based unified networking in the data center. The latest Intel® 10 Gb Ethernet controllers and server adapters include virtualization optimizations and advanced unified networking features. These features have optimizations for lossless Ethernet, intelligent, hardware-based acceleration for (FCoE) and iSCSI, and support for Open FCoE, which is discussed later in this paper.

NetApp provides products and services to plan, design, implement, and maintain Ethernet-based storage solutions.

These solutions simplify network connectivity for today's virtualized servers and lay the foundation for the next generation data center.

The Promise of Ethernet Storage

New usage models and the explosive growth of data in their organizations have forced IT administrators to deal with complicated technical and business challenges. Today, most IT departments deploy separate LAN and storage networks, with storage often divided between network-attached storage (NAS) for file-based applications, and SAN (fibre channel (FC) and iSCSI) for block-based applications. The goal of unified networking is to allow a single-fabric infrastructure, based on 10 GbE, to carry all of these disparate traffic types.

Ethernet has served as a unified data center fabric for years because it supports LAN, NAS, network file system (NFS), common Internet file system (CIFS), and iSCSI SAN traffic. With recent Ethernet enhancements and the ratification of the FCoE specification, data center bridging (DCB)-based Ethernet adapters can now facilitate the connection of servers to FC SANs. The extension of Ethernet's reach and administrators' wide familiarity with FC SAN traffic will help accelerate the move to 10 GbE-based I/O consolidation in virtualized data centers, reduce costs, and improve simplification and agility.

Given its flexibility and long history, it is no surprise that Ethernet storage is the fastest growing segment of the storage system market. The industry research firm IDC*[5] estimates that the worldwide Ethernet-based storage system (NAS and iSCSI SAN) market grew at a compounded annual growth rate (CAGR) of approximately 23 percent between 2005 and 2009. iSCSI storage shipments, in particular, experienced the highest growth rates (70 percent) from 2005-2009, driven by broad iSCSI adoption in Windows*, virtual server, and blade server environments.

Industry analysts project continued gains in the Ethernet storage market share due to increasing deployment of "Ethernet only" data centers (which use a unified 10 GbE infrastructure for all data and storage traffic), the emergence of cloud computing, and the entrance of FCoE solutions into the mainstream.

Ethernet Enhancements for Storage

Data Center Bridging for Lossless Ethernet

To strengthen 10 GbE as a unified data center fabric, the Institute of Electrical and Electronics Engineers (IEEE) has standardized Ethernet enhancements to support storage traffic, including FCoE and iSCSI. These extensions, known collectively as Data Center Bridging, enable better traffic prioritization over a single interface, as well as advanced means to shape traffic on the network to decrease congestion. In short, DCB delivers a lossless Ethernet fabric for storage traffic.

Fibre Channel over Ethernet: Enabling End-to-End Unified I/O

FCoE is a logical extension of Ethernet that uses the fibre channel network, service, and protocol layers to carry data packets over the Ethernet's physical and data link layers. Fibre channel's unique network-centric management model has proven administrative capacities to scale to thousands of end nodes in a data center. The use of fibre channel's upper layers smoothes the transition to FCoE because existing SAN-based applications do not need to change to benefit from the performance and cost benefits of FCoE. The provisioning responsibilities, now split between the server and storage administrators, can be maintained in the transition from fibre channel SANs to Ethernet-based FCoE-powered SANs.

Many enterprises have extensive FC installations, and the availability of native OS FCoE initiators makes the FC SAN easily accessible for any server with a 10 GbE port. Because it allows FC to use DCB-capable Ethernet, FCoE eliminates the need for dedicated FC HBAs, which reduces cabling and switch-port requirements. At the same time, it coexists with established FC infrastructures. The result is a simplified data center infrastructure, lower equipment and power costs, and universal SAN connectivity across the data center over the trusted Ethernet fabric.

Introduction of the Open FCoE Architecture

The Open FCoE approach consists of standard 10 GbE adapters and native operating system-based FCoE initiators, which together provide robust, scalable, and high-performance server connectivity without expensive, proprietary hardware. As shown in Figure 4, Open FCoE implements the complete FC protocol in the operating system kernel. It provides libraries for different system-level implementations, which allow vendors to

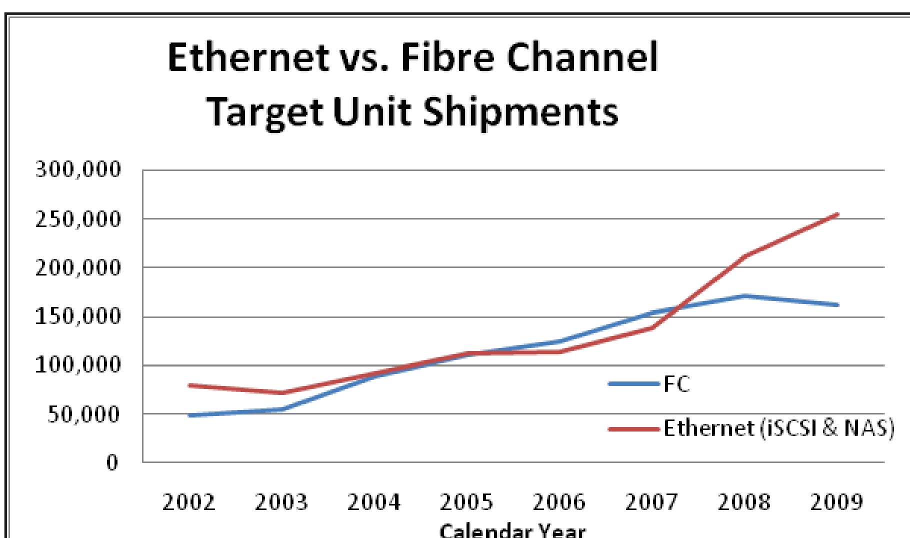


Figure 3. Actual Unit Shipments for Ethernet and Fibre Channel Storage (IDC, 2009)

implement data plan functions of the FCoE stack in the hardware to deliver optimum performance.

Open FCoE: Momentum Started in the Linux* Community

The Linux community accepted the Open FCoE project in November 2007 with a goal to accelerate the development of a native FCoE initiator in the Linux kernel. The industry responded enthusiastically, and today there are over 190 active participants in the community who contribute code, provide review comments, and test the Open FCoE stack. To date, the Open FCoE source website (www.open-fcoe.org) has received over 20,000 hits. Open industry standards and Open Source play a significant role in the modern data center, as they lower R&D investment and enable access to a multi-vendor supply chain that is designed for heterogeneous interoperability. This ultimately results in greater choice and lower equipment costs.

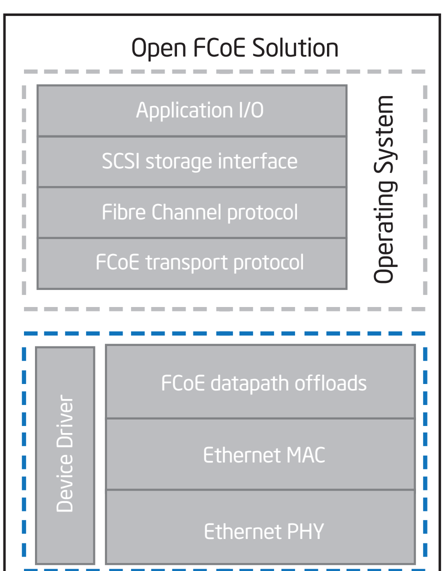


Figure 4. Overview of Open FCoE Initiator Solution with Intel® 10 GbE NIC

Currently, production Open FCoE solutions are available from major vendors such as SUSE* and Microsoft*. Other implementations will be released soon. Contact your OS and virtual machine monitor (VMM) vendors for availability information. We used Microsoft Windows Server* 2008 R2 in our tests for this paper.

The Open FCoE approach offers a number of advantages in terms of accessibility and ease of use.

Accessibility: The Open FCoE approach makes FC SAN access available to any server with a DCB-capable 10 GbE adapter installed. IT departments can standardize on these adapters to simplify server connectivity across multiple use cases. With FCoE support integrated into the operating system, FC SAN access will become readily available across these converged networks and even more accessible once 10 GbE LOM technology becomes more pervasive.

Ease of Use: Because the Open FCoE approach uses DCB-capable 10 GbE adapters that comply with broad, well known Ethernet standards, IT can utilize existing knowledge to configure and manage these adapters for FCoE deployments. In fact, IT can standardize on a single product or product family for all LAN and SAN connectivity. FCoE initiator integration into the OS also means common OS-based tool support across a product family or even adapters from multiple vendors, as they become available.

Native Initiator Success: iSCSI

iSCSI provides an excellent example of the success of native storage initiators integrated into the operating system.

In the early days of iSCSI, proponents of iSCSI HBAs claimed that these dedicated adapters were necessary to deliver acceptable performance. iSCSI HBAs offload iSCSI processing to a separate processor on the adapter, rather than allowing the host processor and operating system to handle these tasks.

Today, all major server operating systems include native iSCSI support, which delivers the same benefits as those detailed above. Native iSCSI initiators have continued to mature and now support advanced adapter features and platform advancements that help deliver excellent performance. In fact, Intel recently demonstrated a standard Intel® Ethernet 10 Gigabit Server Adapter driving 1.25 million input/output operations per second (IOPS) using the native iSCSI initiator in Windows Server 2008 R2.

IT departments which standardize on Intel Ethernet Server Adapters for iSCSI connectivity are able to use a single initiator, TCP/IP stack, and set of management tools and IT policies. Easier server provisioning, lower likelihood of human error, and simpler management enable lower capital and operational expenditures. That standardization also allows IT to enjoy the benefits of simplified management and integrated tools. Intel expects the same benefits from Open FCoE-based solutions.

Test Configuration Overview

Our test configuration is shown in Figure 5. We have two Intel Xeon processor 5600 series-based servers running Microsoft Windows Server 2008 R2. These servers are equipped with Intel® X520 series 10 GbE NICs running native OS initiators for iSCSI and FCoE. The NICs are connected to a Cisco Nexus* 5020 switch with fibre channel support options. This is where the fibre channel naming service runs. The switch is connected with two 10 GbE connections to a NetApp* FAS6280 storage system. This topology provides separate paths for the iSCSI and FCoE traffic, which each run at full

10 GbE speeds from end-to-end. It is also possible to run iSCSI and FCoE on a single connection from the switch to the NetApp storage system.

The hardware and software used to execute the functional and performance testing consisted of:

- Intel® white box dual-processor server
  - Two Intel® Xeon® processor 5680 series-based servers
    - 6 cores / total 12 threads each
    - 3.33GHz clock speed
  - Intel X520 series 10 Gb/s Ethernet adapter
  - 24 GB Memory
  - 2x146 GB HDD RAID 0
  - Microsoft Windows Server 2008 R2
- NetApp FAS6280 storage system
  - Data ONTAP* 8.0.1 operating system
  - X1139 Unified Target Adapter (for FCoE)
  - 2 TB Flash Cache (intelligent read cache)
- Cisco Nexus 5020 unified switch
  - IOS 4.2
- Microsoft Windows Server 2008 R2
  - Native OS FCoE and iSCSI initiators
- Iometer version 2006.07.27 was used for testing throughput
- To provision storage space on the NetApp FAS6280 system, we used NetApp SnapDrive* for Windows* data management software for both iSCSI and FCoE protocols. This simplified the LUN creation and attachment process.

Iometer Configuration Overview

Iometer, by default, will create one worker thread for each processor core/thread detected. In this case, 24 workers were automatically created so that Iometer could fully utilize the Intel Xeon processor architecture.

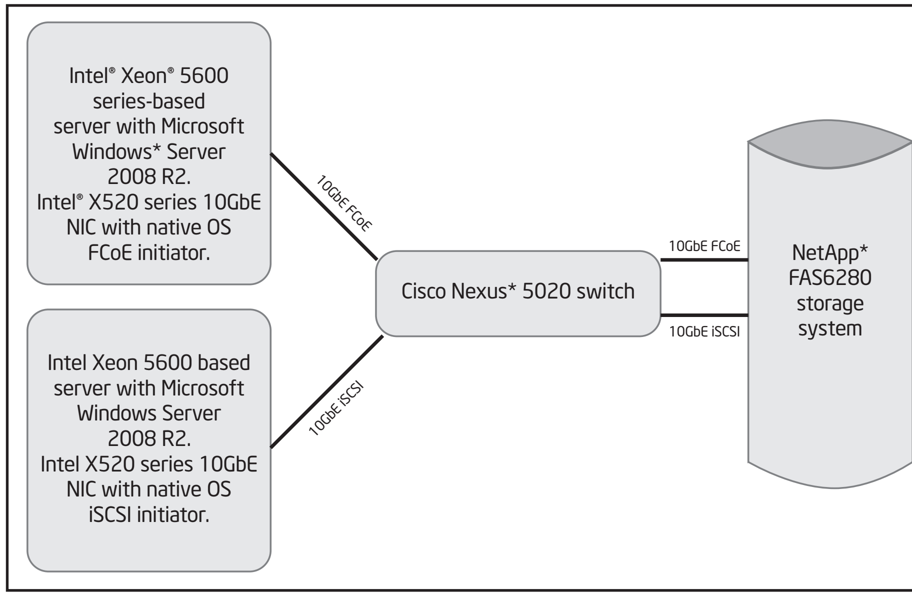


Figure 5

The workers were each configured with disk targets that point to the LUN attached by iSCSI or FCoE protocol, as appropriate.

The number of outstanding I/Os per target was set to eight to increase parallelism and performance.

The access specifications were as follows:

Block sizes used for each FCoE LUN or iSCSI LUN: 4 KB, 8 KB, 16 KB, 32 KB, 64 KB, 128 KB, 256 KB, 512 KB, and 1024 KB.

For each block size we ran seven different tests:

1. 100% sequential read
2. 100% sequential write
3. 100% random read
4. 100% random write
5. 50% read/write sequential mix
6. 50% read/write random mix
7. 50% read/write and 50% random/ sequential mix

Test Configuration Details

NetApp Storage System Configuration Details

NetApp Environment Setup

Hardware

1. NetApp FAS6280 storage system
2. 2.0 TB flash cache memory
3. Three DS14 MK4 shelves with 14 FC 450 GB drives each
4. Two DS4243 shelves with 24 SAS 450 GB drives each
5. One DS4243 shelf with 24 1 TB SATA drives
6. *** Tested LUNs = 100 GB to each server (iSCSI & FCoE)

NetApp Access, Licensing, and Global Configurations

a. Storage is presented to a server as a LUN. LUN creation in NetApp is built within a Volume, which is built within an Aggregate, which consists of one or more redundant array of independent disks (RAID) groups.

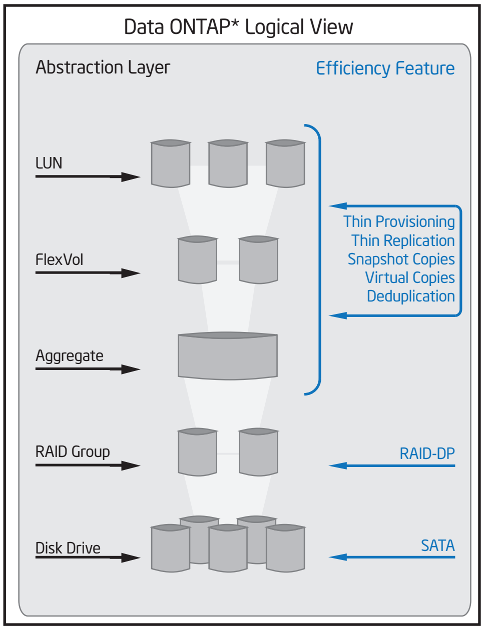


Figure 6

b. Note that we will handle the global configurations in this section. The LUN creation details will be presented in line with the server iSCSI and FCoE configuration sections to follow.

c. Access the storage system through the web portal.

Figure 7

d. Verify that licenses were applied through Manage Licenses in the left menu of the NetApp FilerView* administration tool. Provide the required credentials for VMware vCloud Director to ready the hosts.

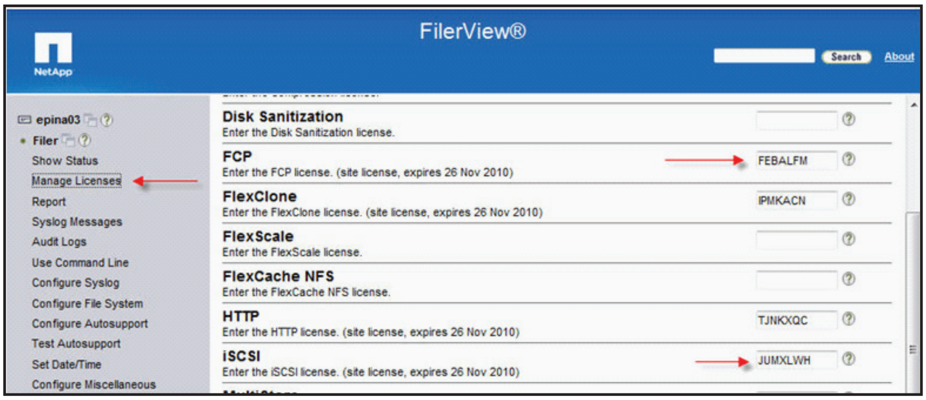


Figure 8

e. NetApp systems use Aggregates to create pools of storage out of one or more RAID groups, which can then be subdivided into logical storage containers as Flexible Volumes. Storage administrators may also take actions such as creating snapshot copies and volume copies at the Aggregate level.

f. To create an Aggregate, open Aggregates from the left menu of FilerView, select "Add" to open the Aggregate Wizard, and select "Next" to start the wizard.

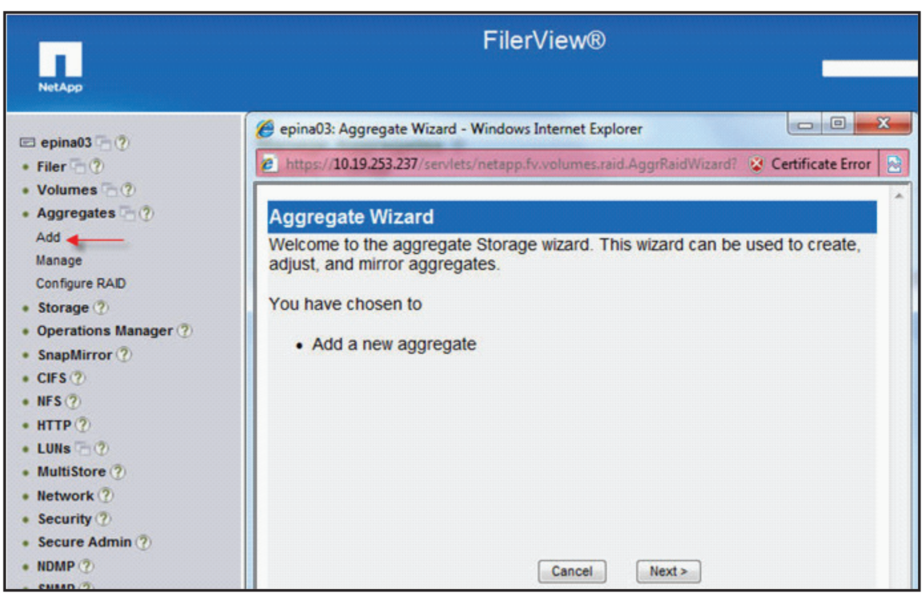


Figure 9

Name the Aggregate and Define its Parameters.

a. Double Parity (DP) is a default. Ensure "Double Parity" is checked to ensure the new aggregate uses a NETAPP RAID-DP* disk layout. RAID-DP further decreases the risk of data loss, as it ensures that parity can never become a performance bottleneck, and that a RAID group can survive the loss of more than one disk.

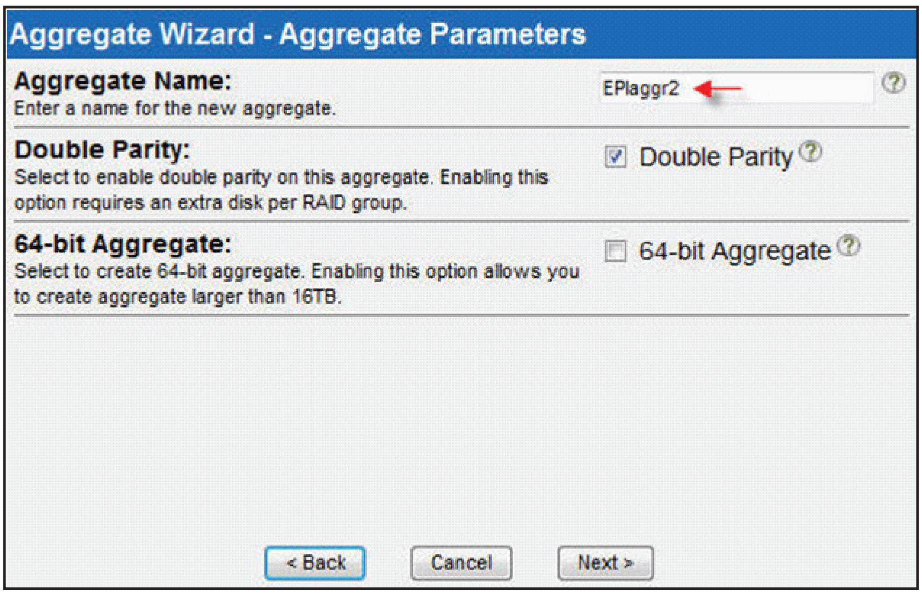


Figure 10

b. Chose the number of disks upon which the RAID group will reside.

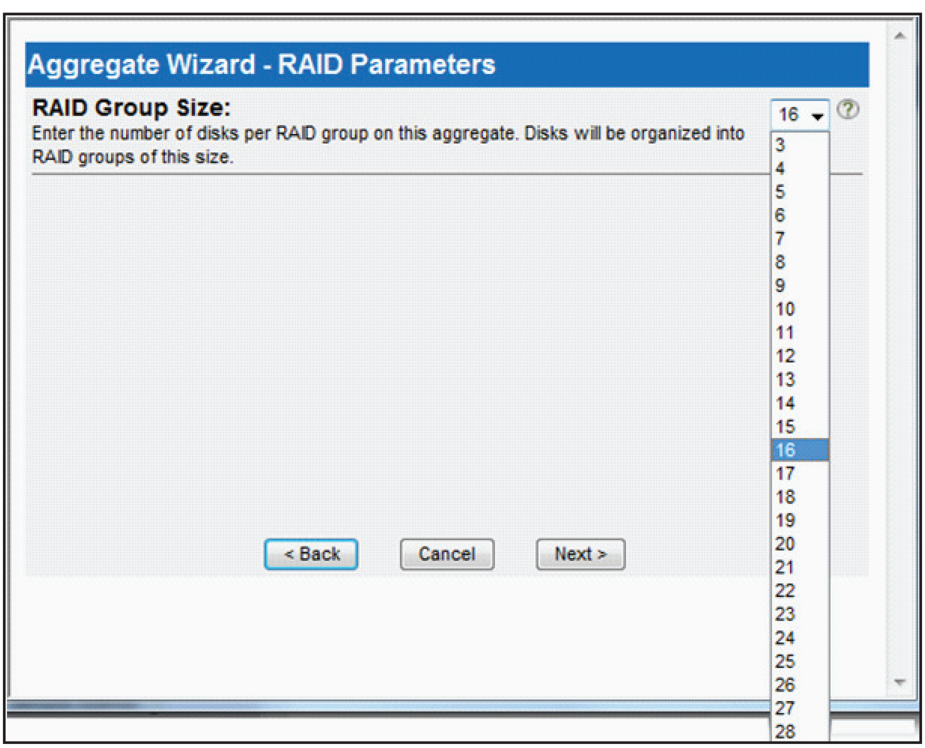


Figure 11

c. NetApp gives the administrator the choice of drives to be used in the Aggregate, and an option to allow the storage system to automatically make the choice.

Figure 12

d. Choose the Disk Type for the Aggregate. The choice is limited to the physical type of installed drives.

Figure 13

e. Select from available drives or let the storage system automatically make the choice.

Figure 14

f. Choose the number of drives for the aggregate.

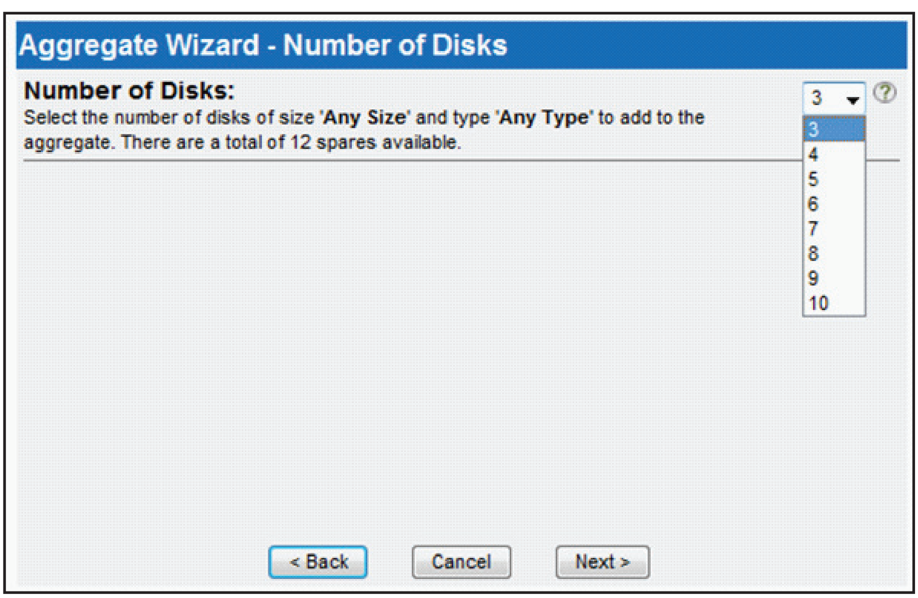


Figure 15

g. Commit the Aggregate to the system.

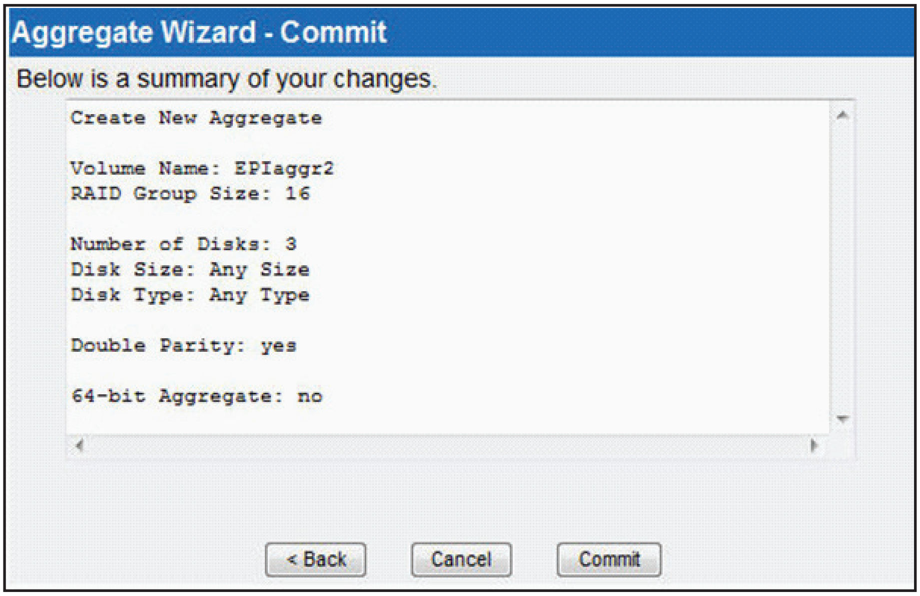


Figure 16

h. Verify the Aggregate creation in the Manage Aggregate window. A refresh of the window may be necessary.

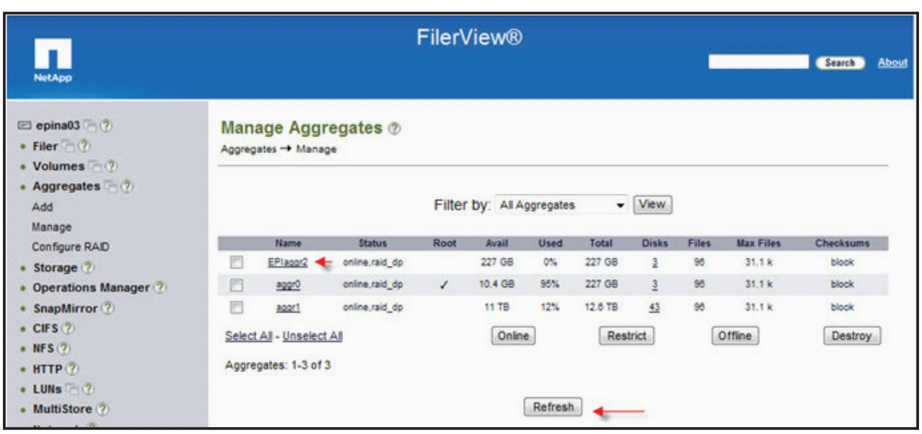


Figure 17

NetApp Creates Volumes, or FlexVols, from the Aggregates

FlexVol volumes are easily created, exported, and mounted onto networked servers.

a. To create a FlexVol volume, open "Volumes" from the left menu of FilerView and select "Add" to open the Volume Wizard.

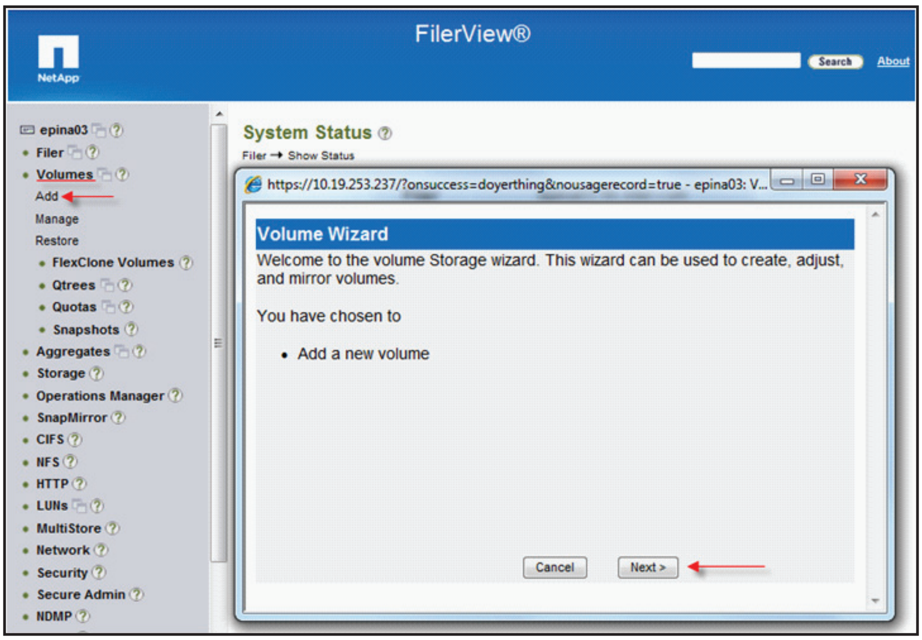


Figure 18

b. Choose "Volume Type" in the Volume Type window.

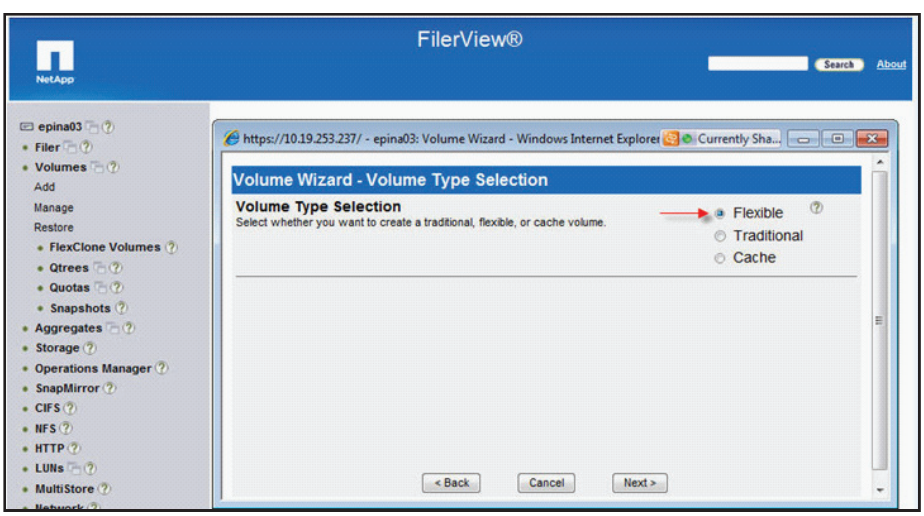


Figure 19

c. Name the volume in the Volume Parameters windows.

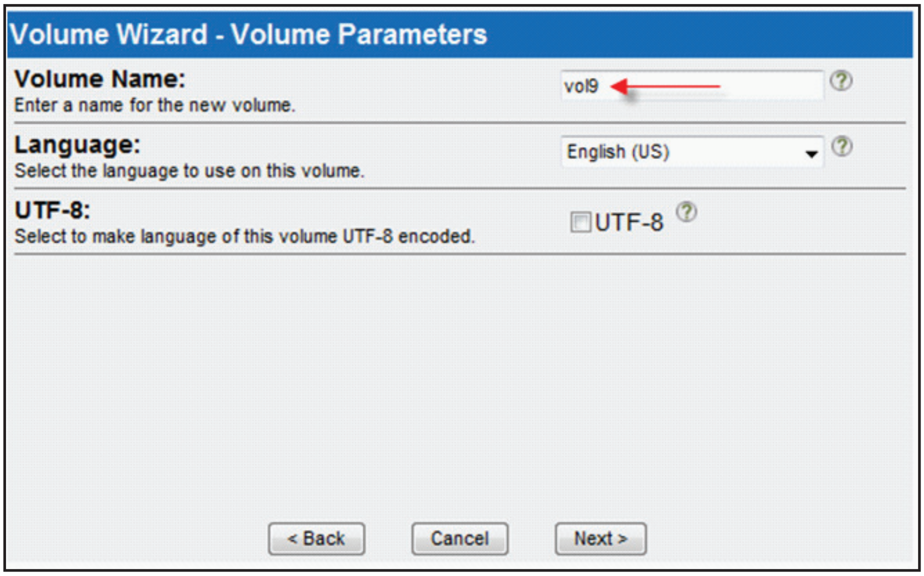


Figure 20

d. Assign the volume to an established Aggregate in the Flexible Volume Parameters window.

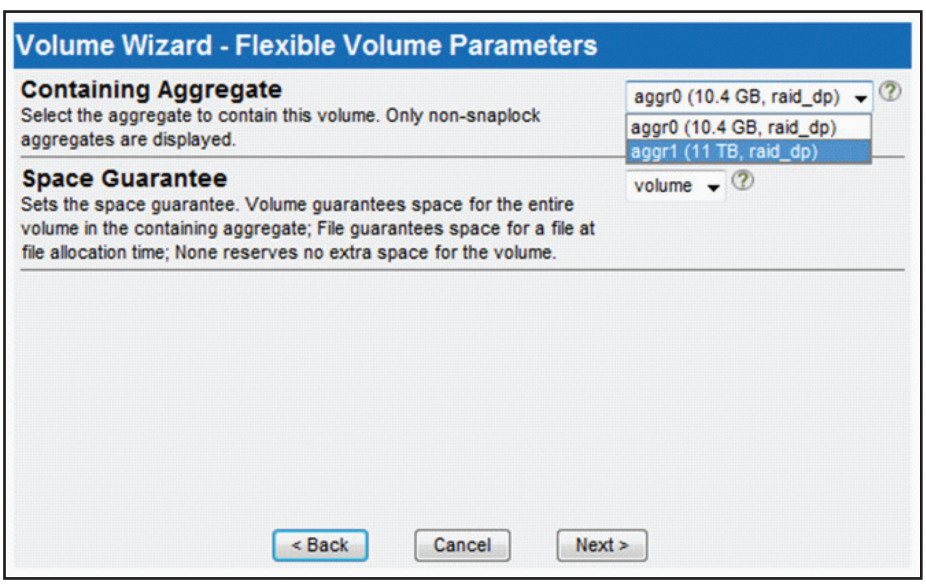


Figure 21

e. Set the volume size in Flexible Volume Size window. "Total Size" will create the volume to the exact specifications set, while "Usable Size" will increase the volume size to enable disk space overhead.

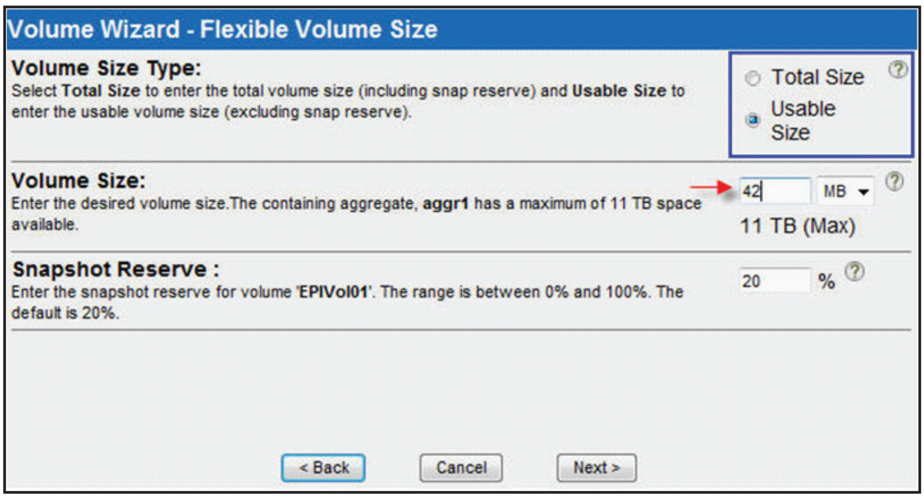


Figure 22

f. Note that we did not need a Snapshot Reserve for our testing. Instead, we just used the default of 20 percent.

g. Commit the volume to the system. ** Note the volume size has increased from the 42 GB initially set due to the Snapshot Reserve.

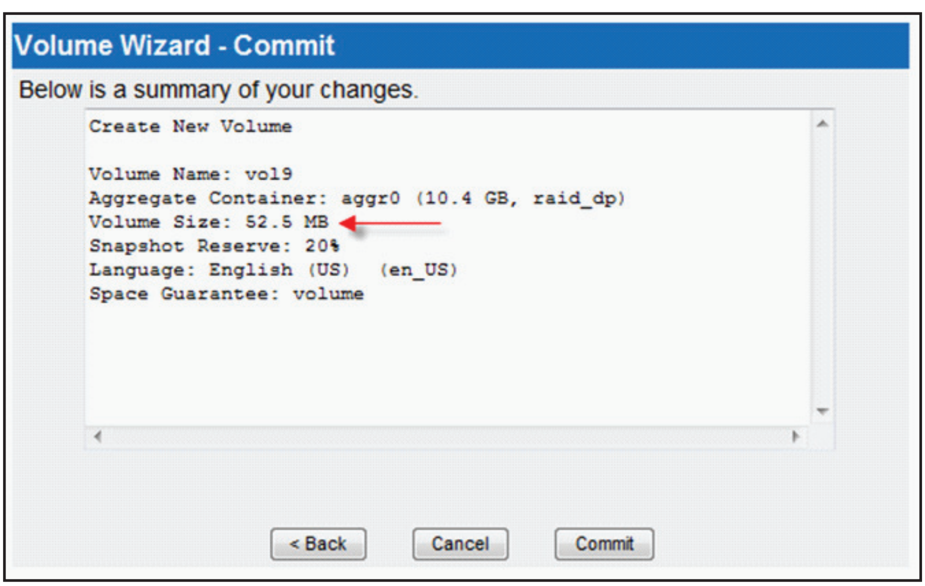


Figure 23

h. Verify the volume in the FilerView Manage Volumes window.

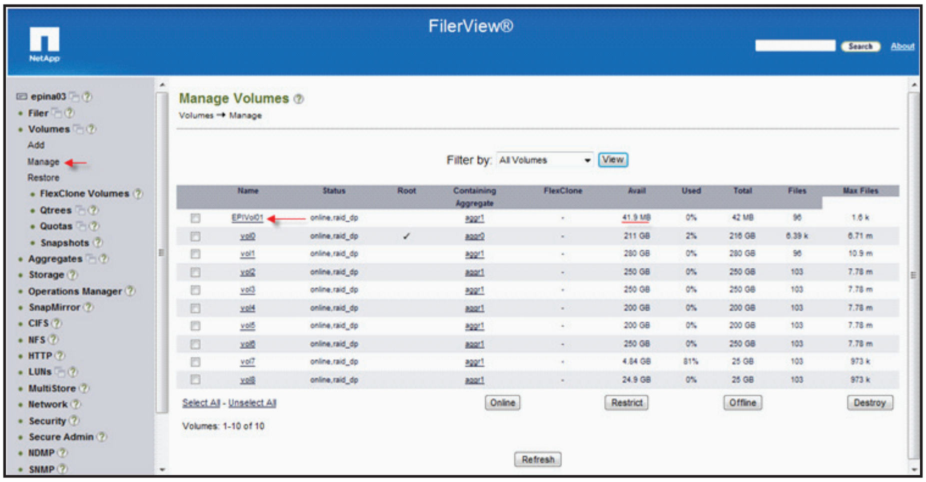


Figure 24

i. Clicking on the volume in the Manage Volumes window will show the configuration parameters of that volume. ** Note the "Total Size" difference from the "Total Capacity" of the volume due to the "Usable Size" setting.

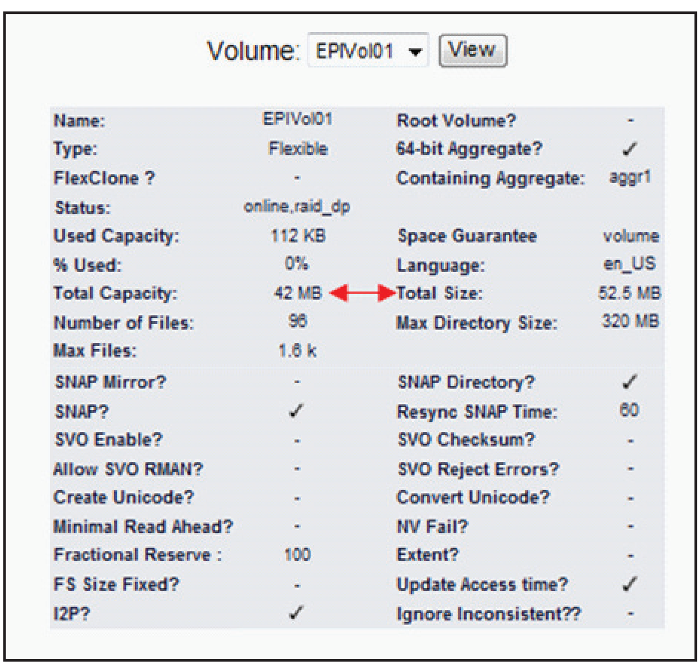


Figure 25

Initiator Groups

Initiator groups are a way for storage administrators to group LUNs according to the parameters of the environments or servers they serve. In our case, this will need to be done for both FCoE and iSCSI initiators.

a. To create an Initiator Group, open from the left menu of FilerView and select LUNs > Initiator Groups > Add to open the Add Initiator Group Window. ** Initiator Groups can be created as part of the LUN creation process.

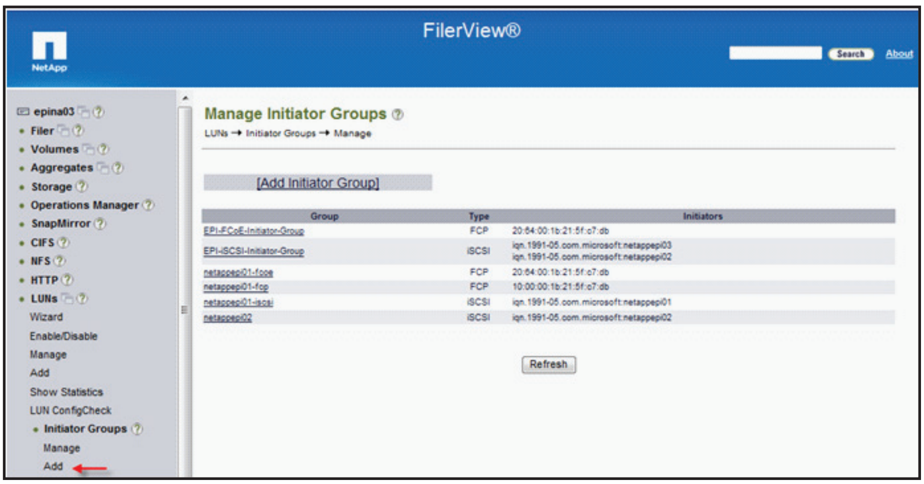


Figure 26

b. Name the Initiator Group. Choose a type (FCP or iSCSI), choose an OS type, and insert the target's worldwide name (WWN) or iSCSI qualified name (IQN) node name. An FCoE Initiator Group is created here.

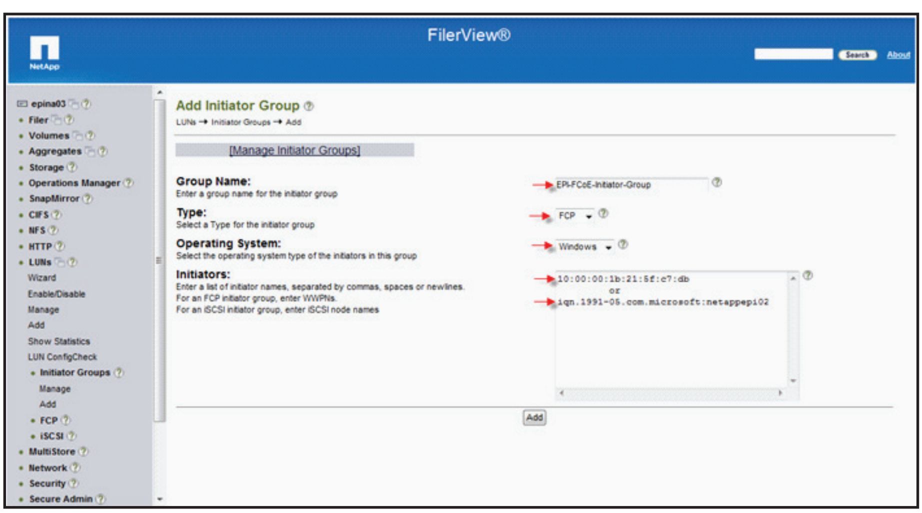


Figure 27

c. Verify the Initiator Group in the Manage Initiator Groups window. A refresh of the window may be necessary.

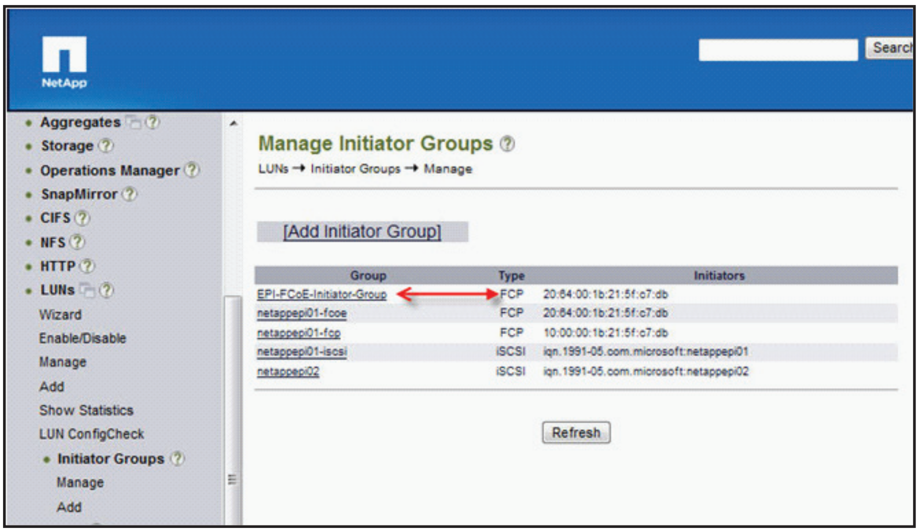


Figure 28

Create LUNs from FlexVol Volumes

A LUN is a logical object with SAN properties.

a. LUNs are managed in the Manage LUNs window found in the left menu of FilerView. Select LUNs > Manage to find the LUNs, LUN description, and associated Initiator Groups.

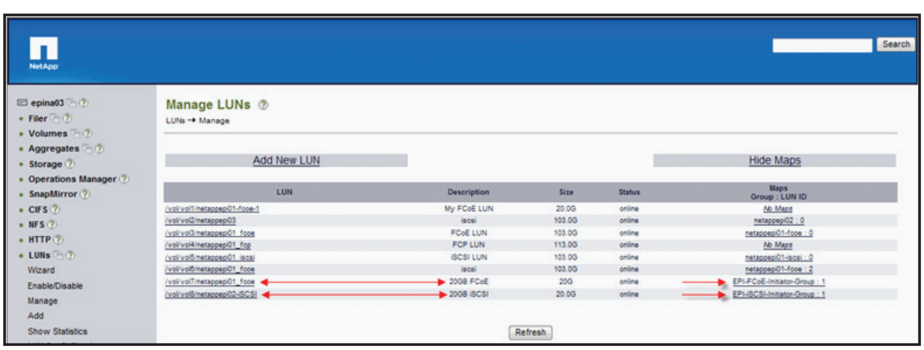


Figure 29

b. We cover the LUN creation details in line with the iSCSI and FCoE server configuration sections to follow.

Cisco Switch Configuration Details

1. Hardware

   a. Cisco Nexus 5020 Chassis 40x10 GE/Supervisor

   b. Intel® Celeron® processor

   c. 2 GB Memory

   d. 4x10 GE + 4x1/2/4G FC Module

2. Software
   a. System Version: 4.2(1)N1(1)
   b. BIOS: version 1.2.0
   c. System image file is: bootflash:/n5000-uk9.4.2.1.N1.1.bin

Detailed Steps

	Command or Action	Purpose
Step 1	Switch# Command terminal	Enters configuration mode.
Step 2	Switch (config)# feature fcoe	Activates FCoE.
Step 3	Switch (config)# policy-map type network-qos jumbo	Configures/creates policy-map of type network-qos with the name jumbo.
Step 4	Switch (config-pmap-nq)# class type network-qos class-default	Sets the class type to class-default (class-default is created on startup by the OS and cannot be deleted).
Step 5	Switch (config-pmap-nq) mtu 9216	Sets frame size to 9216. This is the jumbo frame setting. Jumbo frames provide the best performance for iSCSI with large transfers. Note that FCoE uses jumbo frames by default.
Step 6	Switch (config-pmap-nq)# system qos	Configures qos (Policies are enabled by qos).
Step 7	Switch (config-sys-qos)# service-policy type network-qos jumbo	Configures/creates service-policy of type network-qos with the name jumbo.
Step 8	Switch (config-sys-qos)# vrf context management	Enters VRF context management configuration mode.
Step 9	Switch (config-vrf)# ip route 0.0.0.0/0 10.#.#.#	Configures the IPv4 address of the next hop.
Step 10	Switch (config-vrf)# vlan 41	Creates a VLAN.
Step 11	Switch (config-vlan) name iSCSI	Optional: Name a VLAN.
Step 12	Switch (config-vrf)# vlan 100	Creates a VLAN.
Step 13	Switch (config-vlan) fcoe vsan 2	Enables FCoE for the specified VLAN.
Step 14	Switch (config-vlan) # name FCoE	Optional: Name a VLAN.
Step 15	Switch (config-vlan) # vsan database	Enters VSAN configuration mode.
Step 16	Switch (config-vsan-db) # vsan 2	Adds the VSAN to the VSAN database.
Step 17	Switch (config-vsan-db) # fcdomain fcid database	Enters FC ID database configuration submode.
Step 18	Switch (config-fcid-db) # vsan 2 wwn 20:xx:xx:xx:xx:xx:xx:xx fcid dynamic	Configures a device WWN (20:xx:xx:xx:xx:xx:xx:xx) in the specified VSAN in dynamic mode.
Step 19	Switch (config-fcid-db) # interface vfc01	Creates a virtual fibre channel interface (if it does not already exist) and enters interface configuration mode.
Step 20	Switch (config-if)# bind interface Ethernet1/1	Binds the virtual fibre channel interface to the specified interface.
Step 21	Switch (config-if)# no shutdown	Forces the port state to up.
Step 22	Switch (config-if)# interface vfc02	Creates a virtual fibre channel interface (if it does not already exist) and enters interface configuration mode.
Step 23	Switch (config-if)# bind interface Ethernet1/2	Binds the virtual fibre channel interface to the specified interface.
Step 24	Switch (config-if)# no shutdown	Forces the port state to up.
Step 25	Switch (config-if)# vsan database	Enters VSAN configuration mode.
Step 26	Switch (config-vsan-db)# vsan 2 interface vfc01	Configures the association between the VSAN and virtual fibre channel interface.

Step 27	Switch (config-vsan-db)# vsan 2 interface vfc02	The VSAN number must map to a VLAN on the physical Ethernet interface that is bound to the virtual fibre channel interface.
Step 28	Switch (config-vsan-db)# interface Ethernet1/1	Specifies an interface to configure, and enters interface configuration mode.
Step 29	Switch (config-if)# description FCoE-Server01	Optional: Name a VLAN.
Step 30	Switch (config-if)# switchport mode trunk	Sets an interface as an Ethernet trunk port.
Step 31	Switch (config-if)# switchport trunk allowed vlan 100	Specifies that only certain VLANs are allowed on the specified trunk.
Step 32	Switch (config-if)# spanning-tree port type edge trunk	Explicitly enables edge behavior on the trunk port. Edge ports immediately transition to the forwarding state without passing through the blocking or learning state at linkup.
Step 33	Switch (config-if)# no shut	Forces the port state to up.
Step 34	Switch (config-if)# interface Ethernet1/2	Specifies an interface to configure, and enters interface configuration mode.
Step 35	Switch (config-if)# description Storage 05e	Optional: Name the interface.
Step 36	Switch (config-if)# switchport mode trunk	Sets an interface as an Ethernet trunk port.
Step 37	Switch (config-if)# switchport trunk allowed vlan 100	Specifies that only certain VLANs are allowed on the specified trunk.
Step 38	Switch (config-if)# spanning-tree port type edge trunk	Explicitly enables edge behavior on the trunk port. Edge ports immediately transition to the forwarding state without passing through the blocking or learning state at linkup.
Step 39	Switch (config-if)# no shut	Forces the port state to up.
Step 40	Switch (config-if)# interface Ethernet1/3	Specifies an interface to configure, and enters interface configuration mode.
Step 41	Switch (config-if)# description iSCSI-Server01	Optional: Name the interface.
Step 42	Switch (config-if)# switchport access vlan 41	Sets the access mode of the interface to the specified VLAN.
Step 43	Switch (config-if)# no shut	Forces the port state to up.
Step 44	Switch (config-if)# interface Ethernet1/4	Specifies an interface to configure, and enters interface configuration mode.
Step 45	Switch (config-if)# description Storage 05d	Optional: Name the interface.
Step 46	Switch (config-if)# switchport access vlan 41	Optional: Name the interface.
Step 47	Switch (config-if)# no shut	Forces the port state to up.
Step 48	Switch (config-if)# interface mgmt0	Specifies the management interface to configure, and enters management interface configuration mode.
Step 49	Switch (config-if)# ip address 10.#.#.#	Sets the IP for switch management.
Step 50	Switch (config-if)# zone name zone1 vsan 2	Configures and names zone1 in VSAN2.
Step 51	Switch (config-zone)# member pwwn 20:xx:xx:xx:xx:xx:xx:xx	Adds port world wide name (PWWN) of server's Intel® X520 Series Server Adapter to zone. Perform this step for each server PWWN.
Step 52	Switch (config-zone)# member pwwn 50:xx:xx:xx:xx:xx:xx:xx	Adds port world wide name (PWWN) of NetApp* Unified Target Adapters (UTA) to zone. Perform this step for each member PWWN.
Step 53	Switch (config-zone)# zone name zone1 vsan 2	Re-enters zone set #1.
Step 54	Switch (config -zoneset)# member zone1	Adds a zone as a member of the previously specified zone set.
Step 55	Switch (config -zoneset)# zoneset activate name zoneset1 vsan 2	Activates the zone set and updates the new zone name in the active zone set.

Step 56	Switch (config)# exit	Exits the configuration mode.
Step 57	Switch# copy running-config startup-config	Saves the configuration.
Step 58	Switch# show vlan	Displays selected configuration information for the defined VLAN(s):

```
     VLAN Name                             Status            Ports
---- -------------------------------- --------- -------------------------------------
41   iSCSI                              active             Eth1/2, Eth1/4
100  FCOE                               active             Eth1/1, Eth1/3
```

Step 59	Switch# show vsan 2	Displays information about a specific VSAN:

```
vsan 2 information
         name:VSAN0002 state:active
         interoperability mode:default
         loadbalancing:src-id/dst-id/oxid
         operational state:up
```

Step 60	Switch# show fcoe database	Displays the contents of the FCoE database:

```
-------------------------------------------------------------------------------
INTERFACE       FCID             PORT NAME                MAC ADDRESS
-------------------------------------------------------------------------------
Vfc01          0x880001        20:xx:xx:xx:xx:xx:xx:xx 00:1b:21:5e:c7:eb
Vfc02          0x880004        50:xx:xx:xx:xx:xx:xx:xx 00:c0:de:14:e4:27
```

Step 61	Switch# show interface ethernet 1/1	Displays the physical Ethernet interface; verify if the interface is in trunk mode; verify the traffic flow rate; verify that the jumbo MTU is enabled:

```
Ethernet1/21 is up
  Hardware: 10000 Ethernet, address: 000d.ecb2.8bdc (bia 000d.ecb2.8bdc)
  Description: NetAppEPI01
  MTU 1500 bytes, BW 10000000 Kbit, DLY 10 usec,
     reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation ARPA
  Port mode is trunk
  full-duplex, 10 Gb/s, media type is 10g
  Beacon is turned off
  Input flow-control is off, output flow-control is off
  Rate mode is dedicated
  Switchport monitor is off
  Last link flapped 1d00h
  Last clearing of "show interface" counters 6d00h
  30 seconds input rate 88 bits/sec, 0 packets/sec
  30 seconds output rate 448 bits/sec, 0 packets/sec
  Load-Interval #2: 5 minute (300 seconds)
    input rate 136 bps, 0 pps; output rate 224 bps, 0 pps
```

```
  RX
    6917486410 unicast packets 12947336 multicast packets 2100 broadcast packets
    6930435846 input packets 12891728686220 bytes
    6051166314 jumbo packets 0 storm suppression packets
    0 runts 0 giants 0 CRC 0 no buffer
    0 input error 0 short frame 0 overrun 0 underrun 0 ignored
    0 watchdog 0 bad etype drop 0 bad proto drop 0 if down drop
    0 input with dribble 0 input discard
    12913702 Rx pause
  TX
    7594352404 unicast packets 1139867 multicast packets 2028 broadcast packets
    7595494299 output packets 13335728311942 bytes
    6075073811 jumbo packets
    0 output errors 0 collision 0 deferred 0 late collision
    0 lost carrier 0 no carrier 0 babble
    711978 Tx pause
  5 interface resets
```

Step 63	Switch# Show interface ethernet 1/1 fcoe	Displays the FCoE settings for an interface:

```
Ethernet1/1 is FCoE UP
    vfc01 is Up
        FCID is 0x880003
        PWWN is 20:xx:xx:xx:xx:xx:xx:x
        MAC addr is 00:1c:21:5f:c7:eb
```

Step 64	Switch# show vlan fcoe	Displays the mapping of FCoE VLANs to VSANs:

```
VLAN                VSAN                Status
--------  --------  --------  --------  -------------
100                 2                   Operational
```

iSCSI Configuration Details

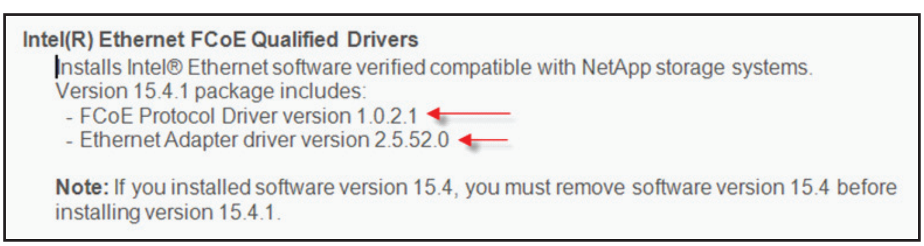
To configure the iSCSI connection and maximize TCP/IP throughput, the following networking best practices were observed:

- Receive Side Scaling was enabled on the host's Intel X520 Ethernet adapter.
  - This feature maximizes the large number of available processor cores in systems that use the Intel Xeon processors because it distributes the work of processing traffic among multiple processors.
  - Note: this feature is enabled by default in Windows Server 2008 R2.
  - See Step 3d below in this section to see how to verify this setting.
- Large Segment Offload was also enabled on the host's X520 Ethernet adapter.
  - This feature moves the work of splitting large data requests into segments small enough to be transmitted by the IP protocol to the Ethernet adapter.
  - Note: this feature is enabled by default in Windows Server 2008 R2.
  - See Step 3e below in this section to see how to verify this setting.
- Jumbo frames of 9000 bytes were configured on the host, switch, and storage.
  - This reduces the amount of network overhead required for TCP/IP headers and checksums as it decreases the ratio of metadata to data in each packet sent or received over the standard MTU size of 1500 bytes.

  - See Step 3c below in this section to see the server setup for jumbo frames.
  - See Step 5 in the Cisco switch configuration above for jumbo frame setup.
  - See Step 6a below in this section for the NetApp storage system jumbo frame setup.
- Multiple TCP streams were established between host and storage through use of a multi-connection session (MCS).
  - MCS allows multiple TCP streams to carry iSCSI traffic. This feature improves performance because it adds parallelism.
  - See Step 10 below in this section to see the setup for MCS.
  - Note: The use of the Windows 2008 R2 standard Microsoft* Multipath I/O (MPIO) implementation is another way to take advantage of the benefits of multiple TCP streams.

Server

1. Hardware
   a. Intel Xeon processor 5680 series @ 3.33 GHz
   b. 24 GB Memory
   c. 2x146 GB HDD RAID 0
   d. Intel® Niantic NIC driver update 2.5.52.0
2. Software
   a. Windows Server 2008 R2
   b. Install the latest Intel X520 10 GB driver update
      i. Find the driver at http://downloadcenter.intel.com
      ii. Search by:
         1. Product Family = Network Connectivity
         2. Product Line = Intel Server Adapters
         3. Product Name = Intel Ethernet Server Adapter X520 Series
            a. The current qualified FCoE Drivers are in Package Version 15.4.1.
         4. We used the Ethernet Driver version 2.5.52.0



Intel(R) Ethernet FCoE Qualified Drivers
Installs Intel® Ethernet software verified compatible with NetApp storage systems.
Version 15.4.1 package includes:
- FCoE Protocol Driver version 1.0.2.1
- Ethernet Adapter driver version 2.5.52.0

Note: If you installed software version 15.4, you must remove software version 15.4 before installing version 15.4.1.

Figure 30

a. Windows Updates
b. Firewall & IE ESC both turned off
c. Jumbo frames enabled

3. iSCSI NIC Properties Setup

a. Launch the Server Manager utility and browse to Network Adapters. Right click and open "Properties."

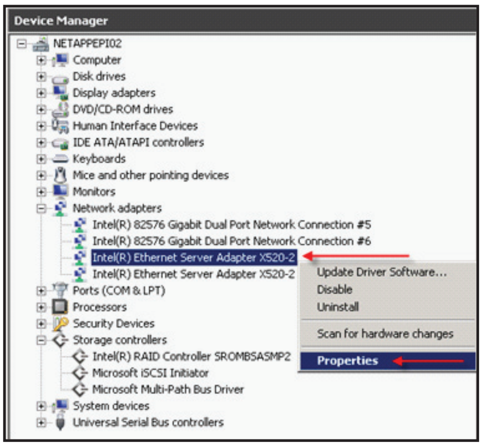


Figure 31

b. Access the "Driver" tab to verify driver.

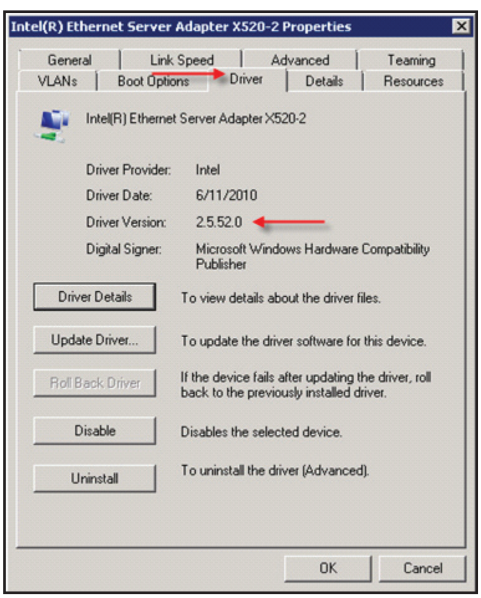


Figure 32

c. Access the "Advanced" tab to make NIC setting changes. We enabled jumbo frames here.

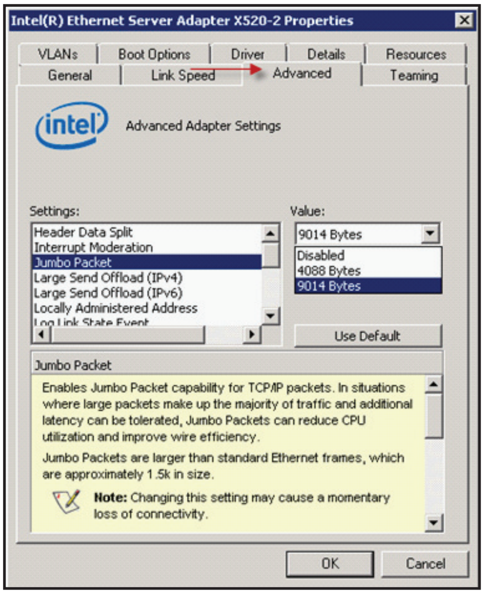


Figure 33

d. Also on the "Advanced" tab, check the setting for Receive Side Scaling. It should be enabled by default, but double check.

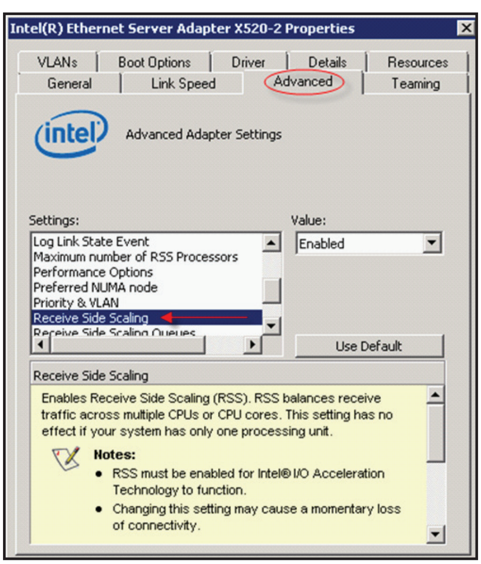


Figure 34

e. Also on the "Advanced" tab, check the setting for Large Send Offload. It should be enabled by default, but double check.

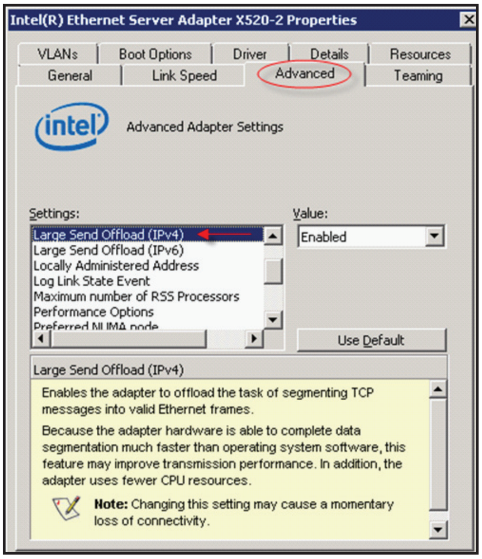


Figure 35

4. Server iSCSI Initiator

a. From the server's Start menu, navigate to Administrative Tools > iSCSI Initiator.

b. After launching, go to the "Configuration" tab to get the Initiator Name. The Initiator Name will be used on the storage device to create a target.

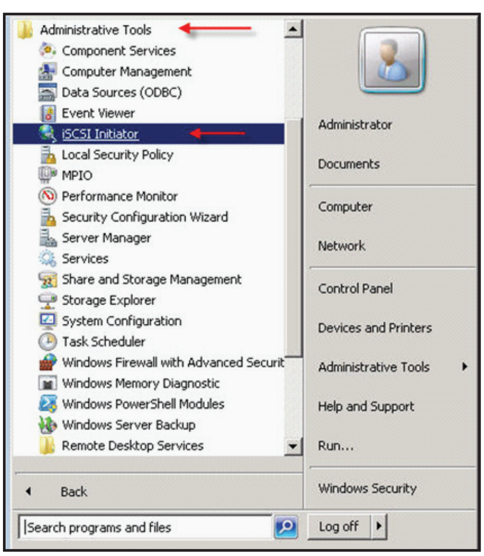


Figure 36

c. If this is the first time the iSCSI Initiator has been run, the system will report that the iSCSI service is not running and will ask for permission to start the service. This only happens the first time.

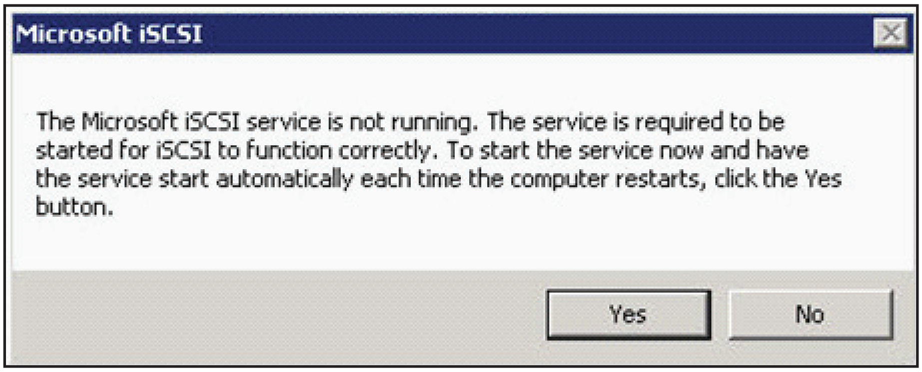


Figure 37

d. After launch, navigate to the “Configuration” tab of the iSCSI initiator properties and capture the server’s initiator name.

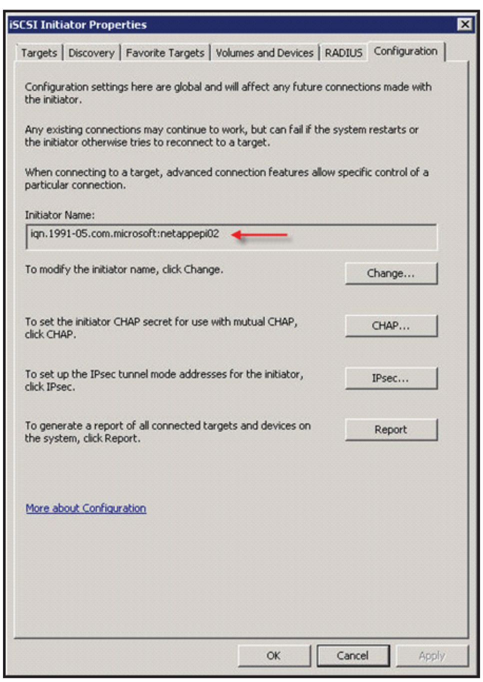


Figure 38

5. **Stop!** Proceed to the NetApp storage system to make the necessary changes to configure and assign a storage target to the server.

6. iSCSI Configuration on the NetApp storage system:

   a. A best practice for iSCSI is to use jumbo frames. This provides the best performance for larger transfer sizes.

      i. From the command line type the following command:

         1. “ifconfig <interface> mtusize 9000”
         2. Repeat the above command line for each interface that is used for iSCSI protocol.

      ii. You can verify the jumbo frame setting through the GUI. See Appendix B for details.

b. iSCSI is enabled by default on all of the ports of the NetApp storage system. As a best practice, it should be turned off on ports that do not require iSCSI usage. To turn it off, select "Manage Interfaces" under LUNs > iSCSI from the left menu of FilerView.

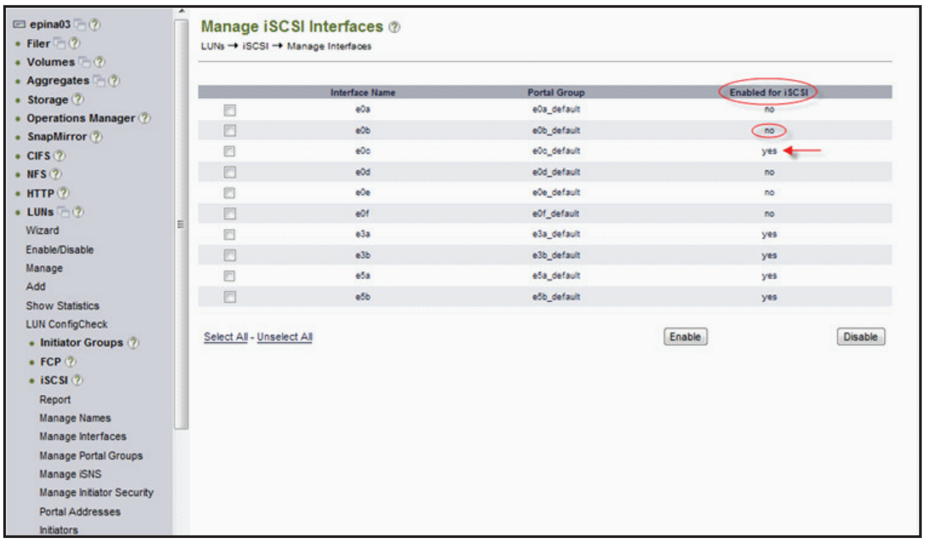


Figure 39

c. To create an iSCSI LUN, open LUNs from the left menu of FilerView and select "Wizard" to open the LUN Wizard.

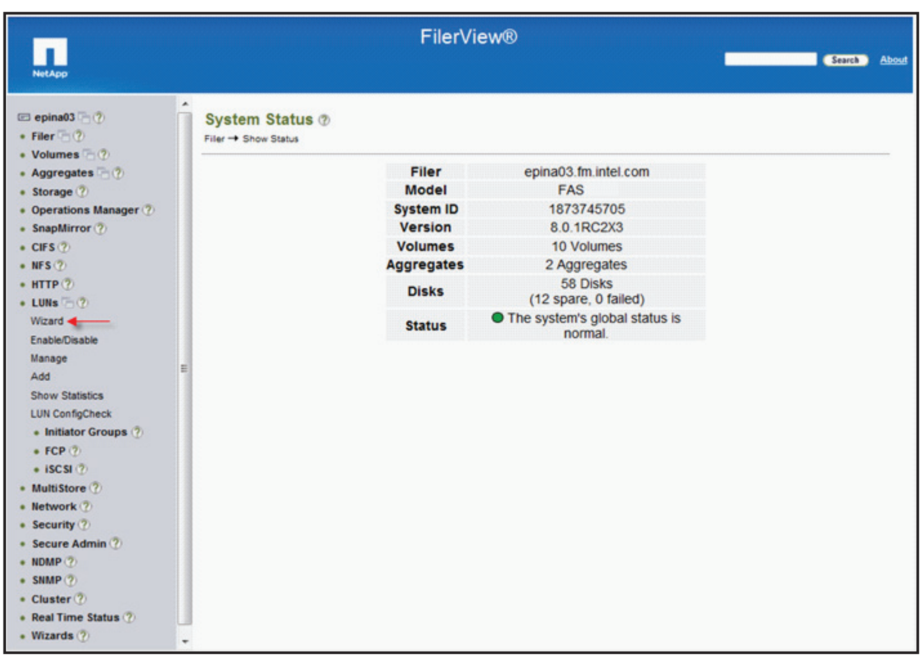


Figure 40

d. Select “Next” at the LUN Wizard Welcome screen.

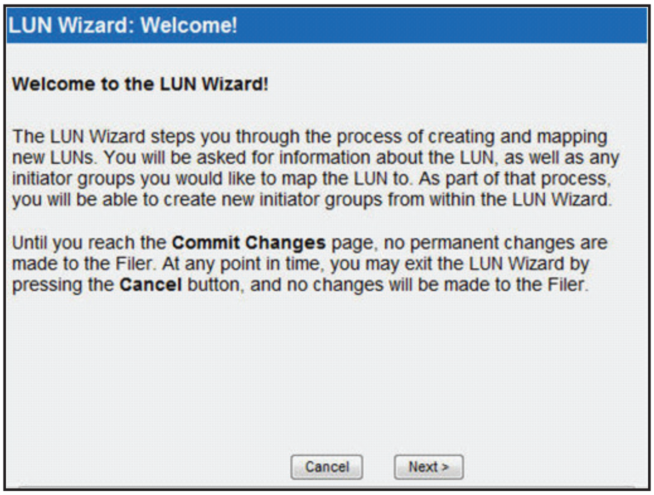


Figure 41

e. Set LUN configurations at the Specify LUN Parameters window. Path information is available through the Filer Manage Volumes window; the name is administrator defined. Set LUN size, protocol type, and select “Next.”

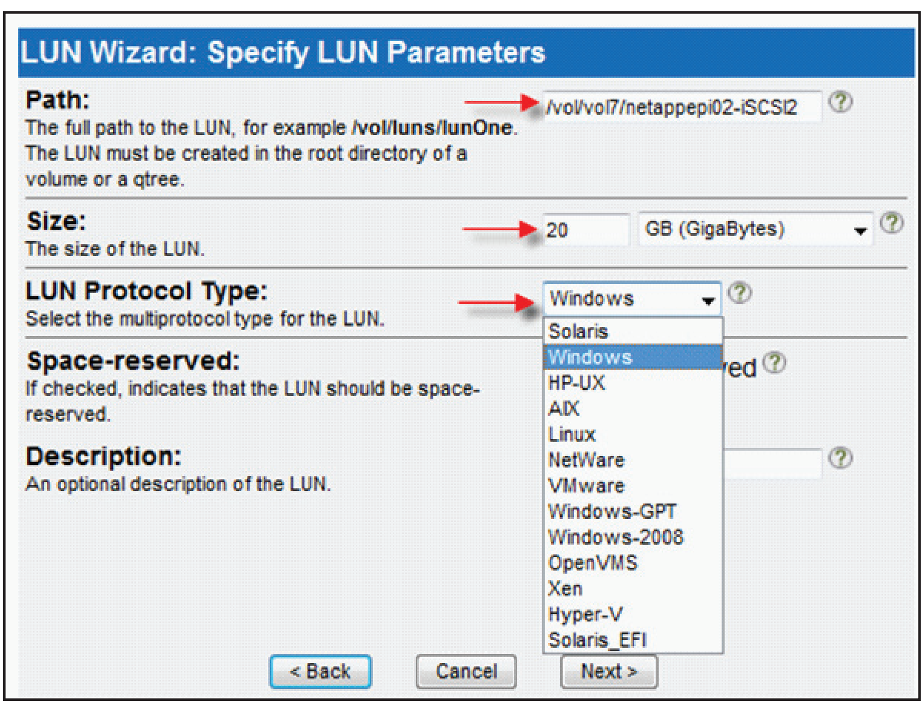


Figure 42

f. For a LUN to be activated, it must be added to an Initiator Group. Initiator Groups can be created during LUN creation or independently. This study used both methods: independently in FCoE and during LUN creation with iSCSI. Select "Add Group" to configure LUN with an existing Initiator Group.

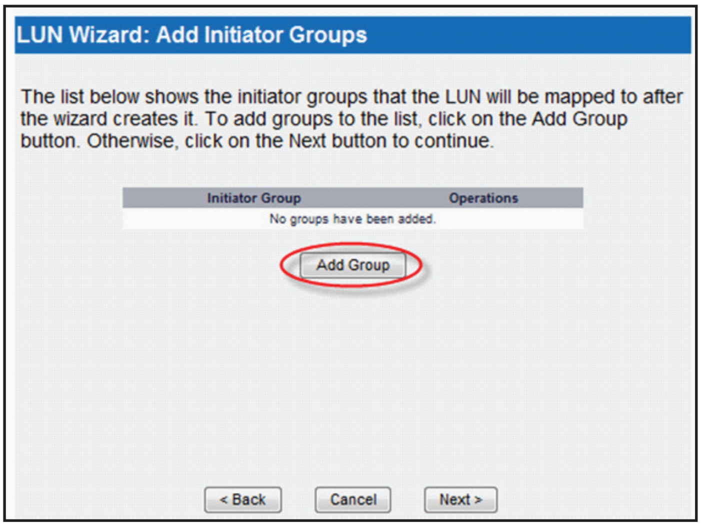


Figure 43

g. Select "Use existing initiator group" and select "Next" to open the LUN Map Add Group window.

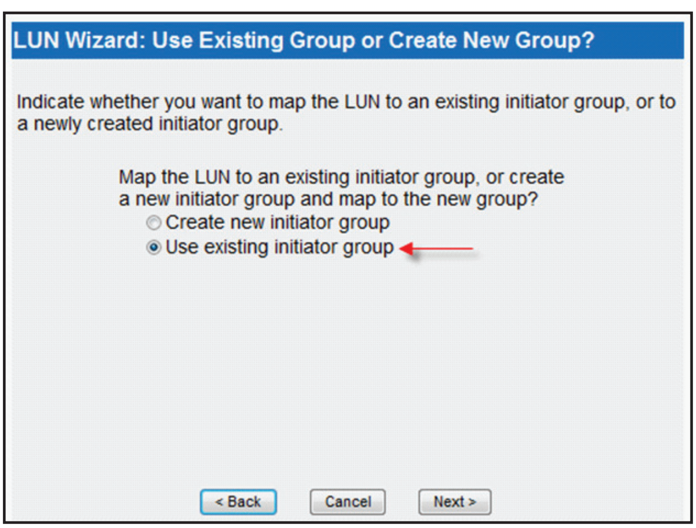


Figure 44

h. The LUN Map Add Groups window opens. Select the "Existing iSCSI Initiator Group" and select "Next."

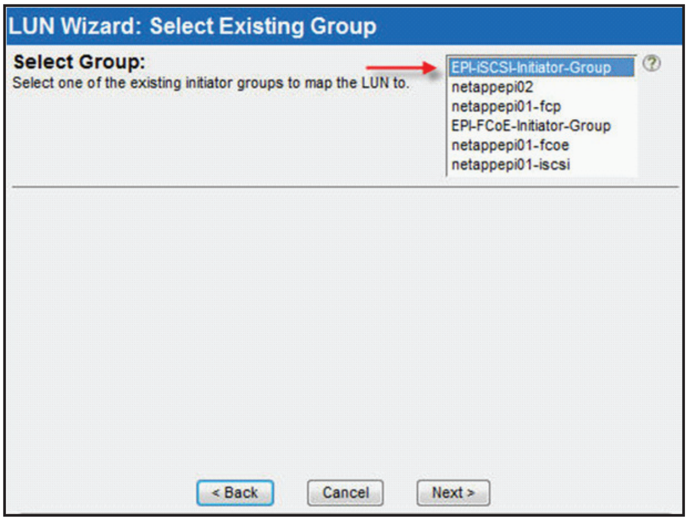


Figure 45

i. Select "Next" at the Add Initiator Groups window.

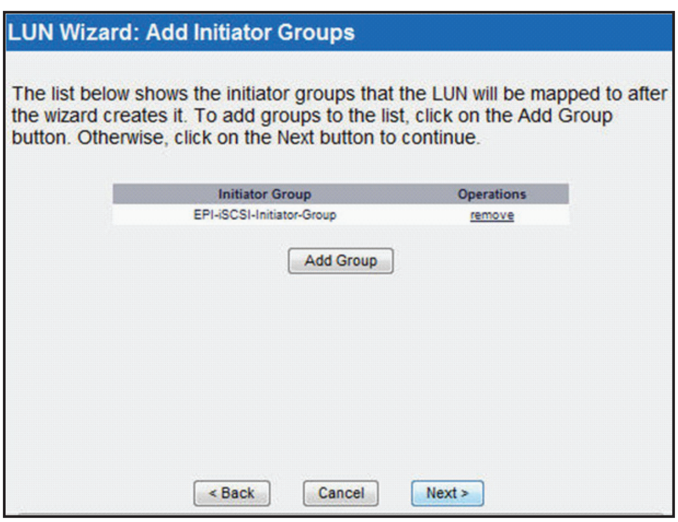


Figure 46

j. Assign a LUN ID number. If left blank, the storage system will assign one automatically.

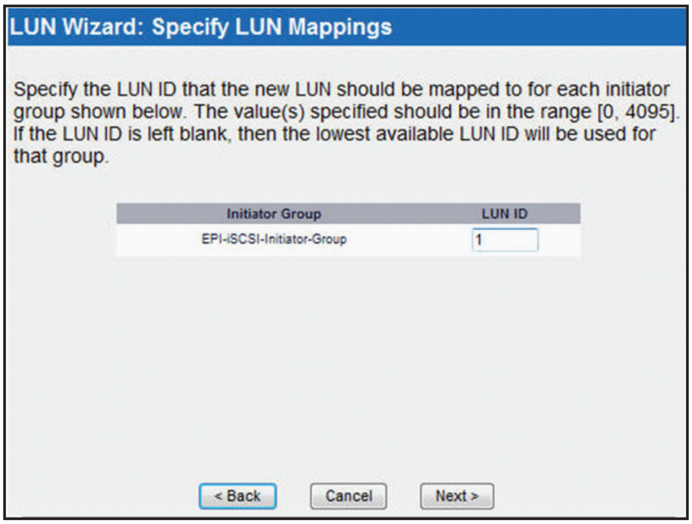


Figure 47

k. Select “Commit” to save the changes made in the LUN Wizard.

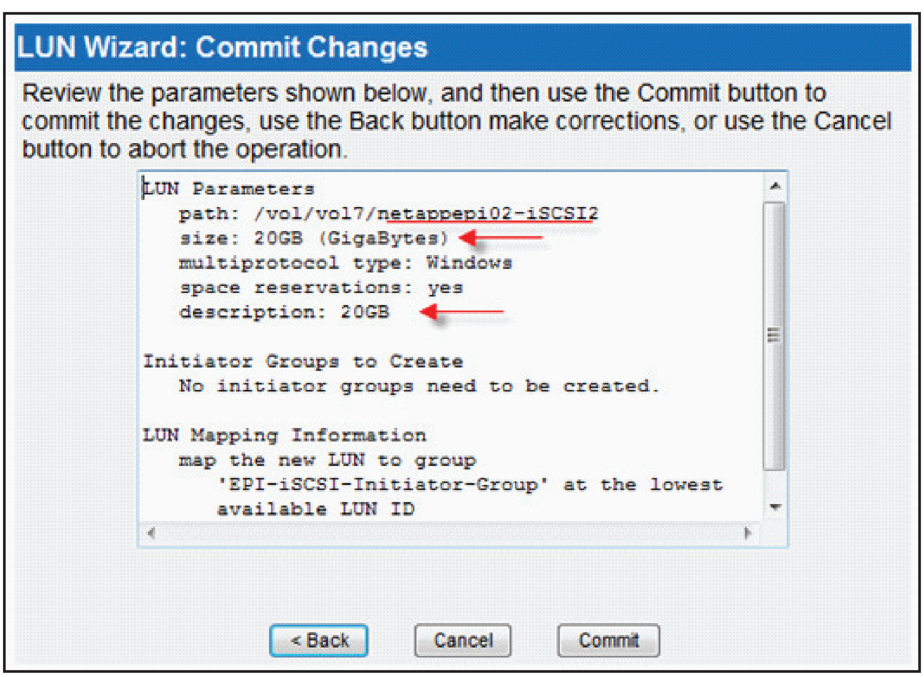


Figure 48

l. Select “Close Window” to complete the changes.

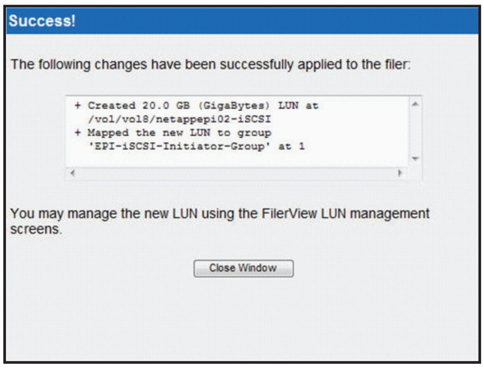


Figure 49

m.Verify the LUN’s creation in the Manage LUNs window. A refresh may be necessary.

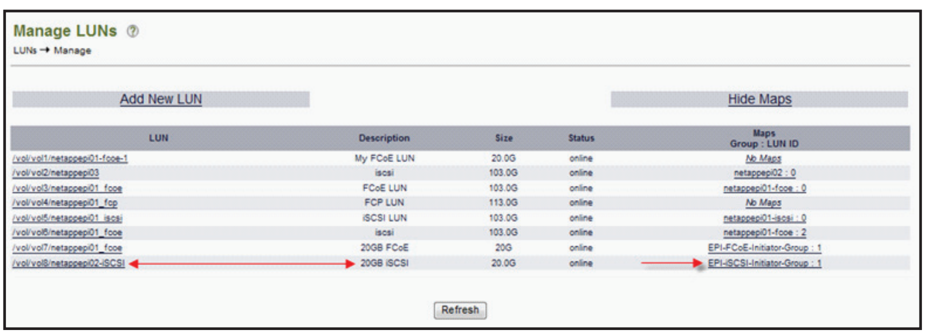


Figure 50

7. **Stop!** The LUN is now created, so move back to the server and add the imported drive.

8. iSCSI Target Discovery and Final Setup

a. Upon completion of the target setup and assignment, re-launch the iSCSI Initiator.

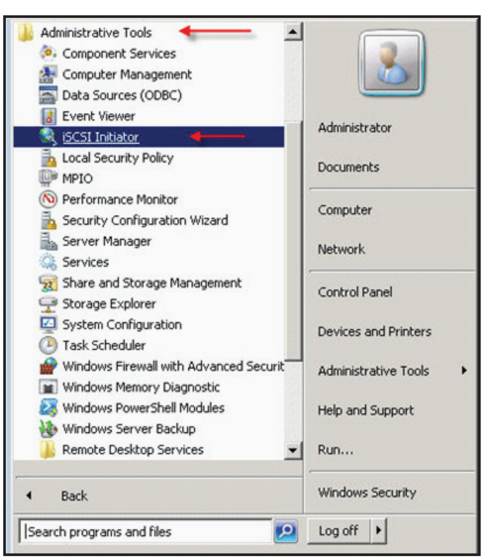


Figure 51

b. Navigate to the “Discovery” tab of the iSCSI Initiator properties window and launch the Discover Portal.

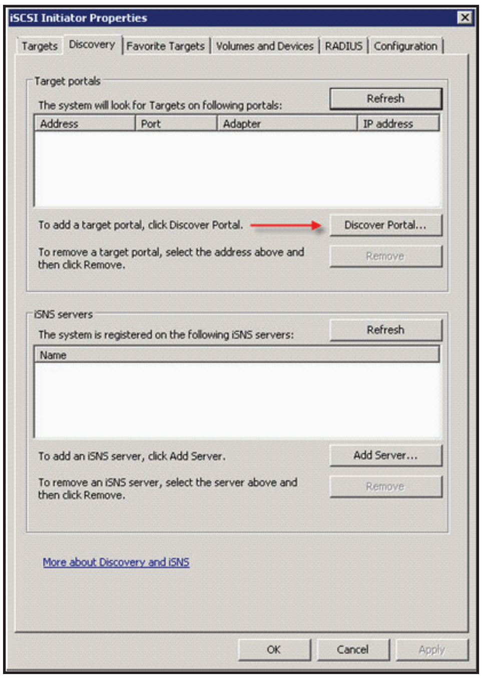


Figure 52

c. In the Discover Target Portal, enter the IP address for the NetApp iSCSI interface and press the “OK” button.

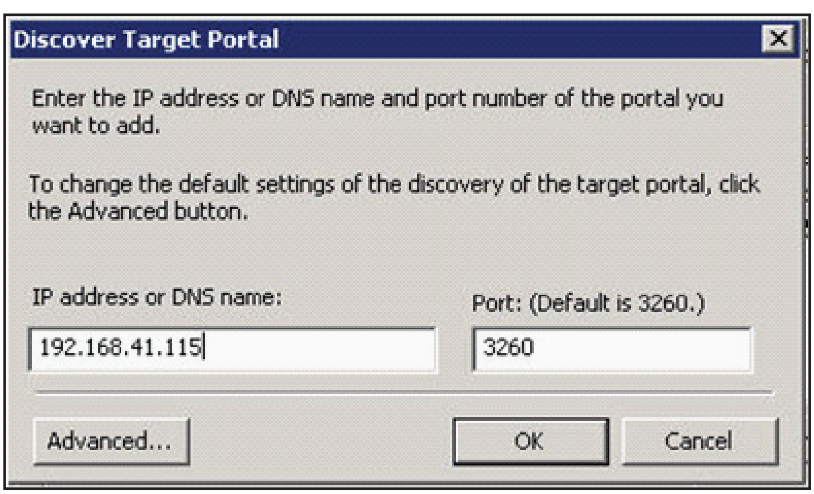


Figure 53

d. Navigate to the "Targets" tab of the iSCSI Initiator Properties window and press the "Refresh" button. The iSCSI Target should now be present in the Discovered Targets window.

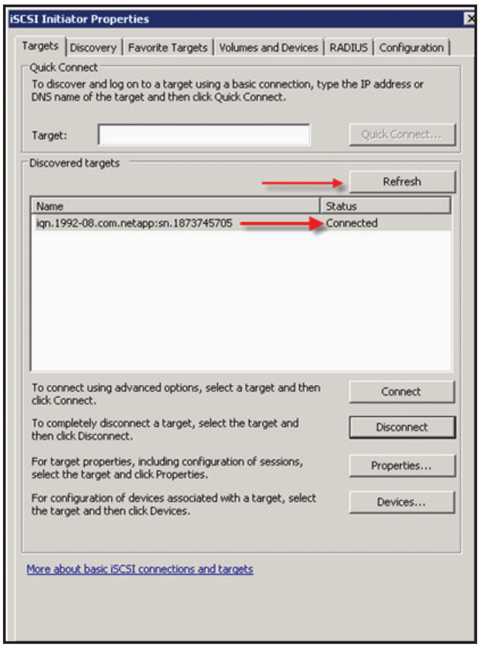


Figure 54

9. Initializing Drives: If network and storage parameters are correctly configured, LUNs will be applied to the server. These drives will act like hot-swapped, imported drives which need to be brought online, initialized, and formatted for use.

a. Enter the Server Manager utility (right-click "Computer" and select "Manage") and then the Storage > Disk Management tool.

b. Right click the unknown disk and bring it "Online."

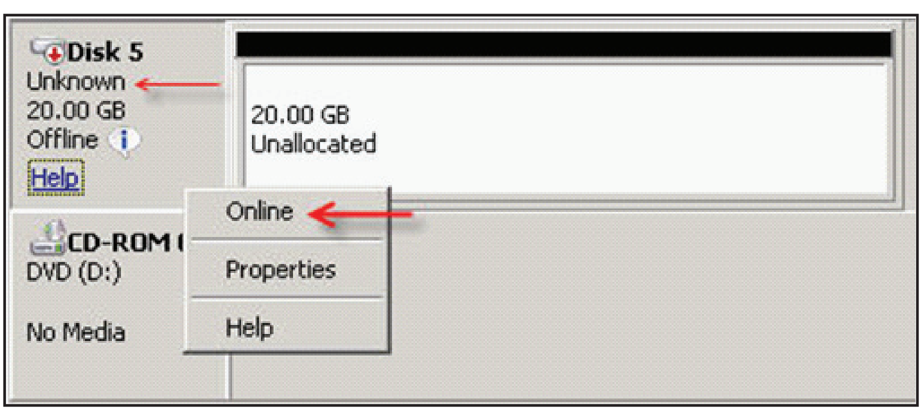


Figure 55

c. Right click the unknown disk to initialize the disk.

Figure 56

d. Create a new simple volume and format for use.

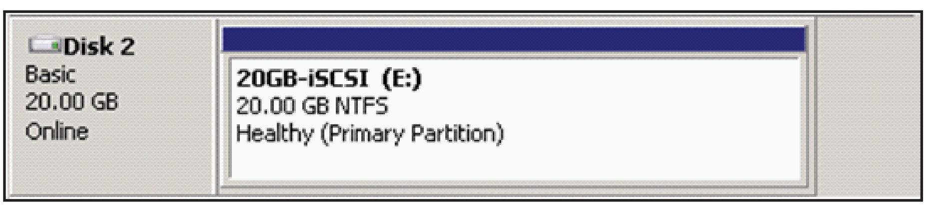


Figure 57

e. Addition of multiple connection sessions (MCS). This feature improves performance because it adds parallelism.[6]

f. Re-launch the iSCSI Initiator.

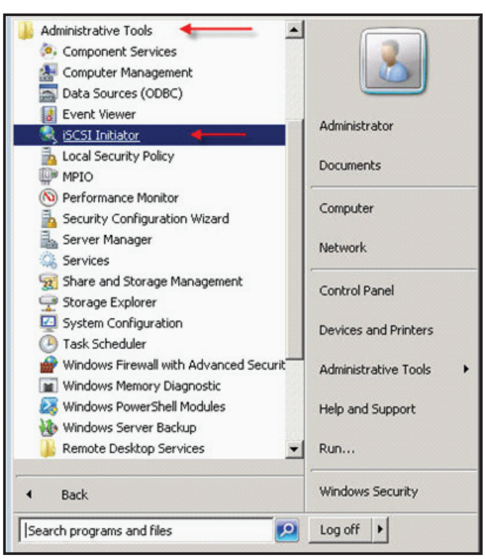


Figure 58

g. Navigate to the "Targets" tab of the iSCSI Initiator properties window and launch the Discover Portal. Then highlight and select the target's properties.

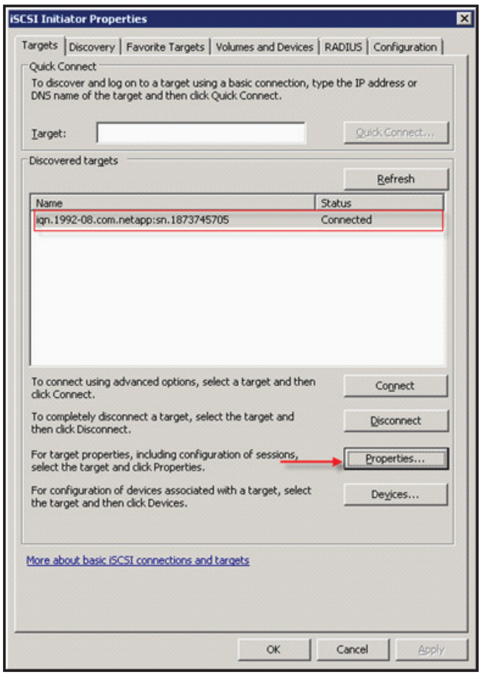


Figure 59

h. Select the "Multiple Connected Sessions" (MCS).

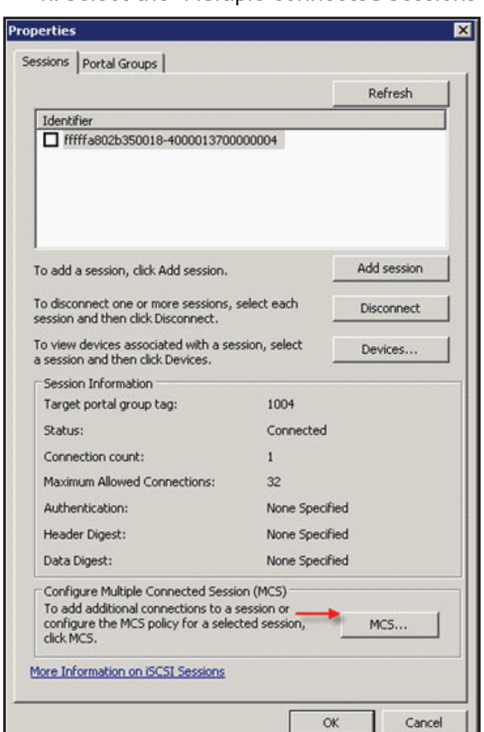


Figure 60

i. Select the "Add" button to allocate another test connection.

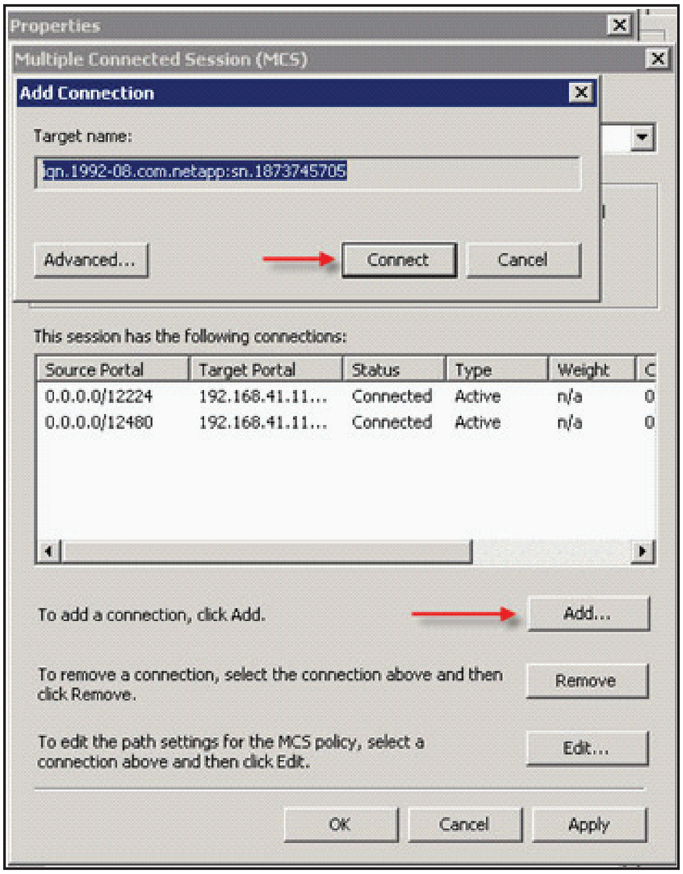


Figure 61

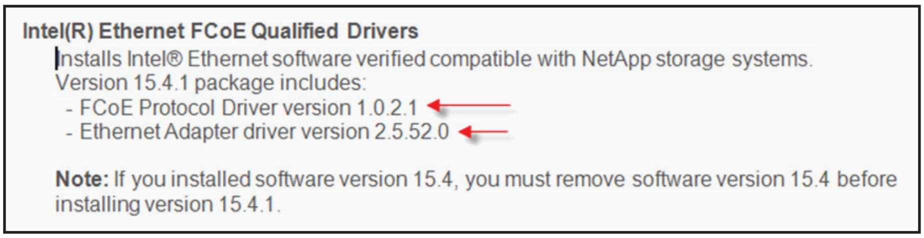
j. In our testing, we used eight sessions.

FCoE Configuration Details

FCoE Server Setup

1. Hardware
   a. Intel Xeon 5680 3.33 GHz
   b. 24 GB Memory
   c. 2x146 GB HDD RAID 0
   d. Intel X520 Niantic NIC driver update 2.5.52.0
2. Software
   a. Windows Server 2008 R2
   b. Install the latest Intel X520 10 GB driver update
      i. Find the driver at http://downloadcenter.intel.com
      ii. Search by:
         1. Product Family = Network Connectivity
         2. Product Line = Intel Server Adapters

3. Product Name = Intel® Ethernet Server Adapter X520 Series
    a. The current qualified FCoE Drivers are in Package Version 15.4.1
        i. We used the Ethernet Driver version 2.5.52.0



Intel(R) Ethernet FCoE Qualified Drivers

Installs Intel® Ethernet software verified compatible with NetApp storage systems.
Version 15.4.1 package includes:
- FCoE Protocol Driver version 1.0.2.1
- Ethernet Adapter driver version 2.5.52.0

Note: If you installed software version 15.4, you must remove software version 15.4 before installing version 15.4.1.

Figure 62

c. Windows Updates

d. Firewall & IE ESC both turned off

3. Intel X520 NIC Installation with FCoE Option

a. Launch the Intel X520 (PROWinx64) 10 GB driver update and accept the Licensing Agreements.

i. At the "Setup Options" screen, check the "Data Center Bridging/Fibre Channel over Ethernet" option and complete the driver update.

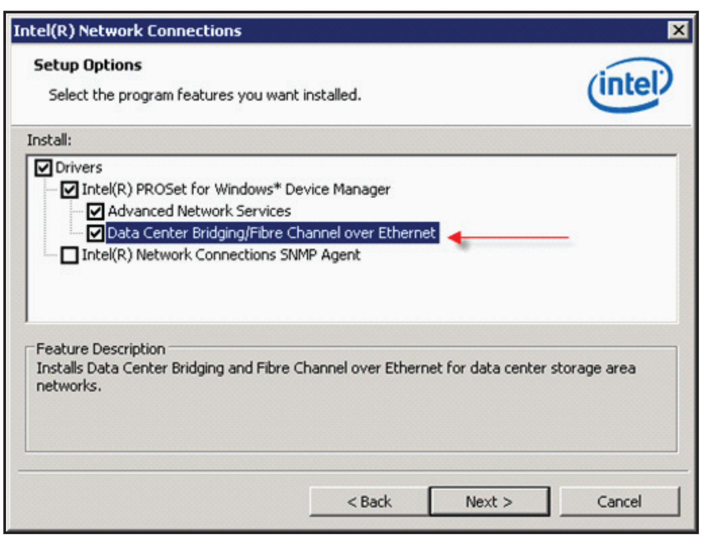


Figure 63

4. FCoE Discovery

a. Launch the Server Manager utility and browse to "Network Adapters." Right click and open "Properties."

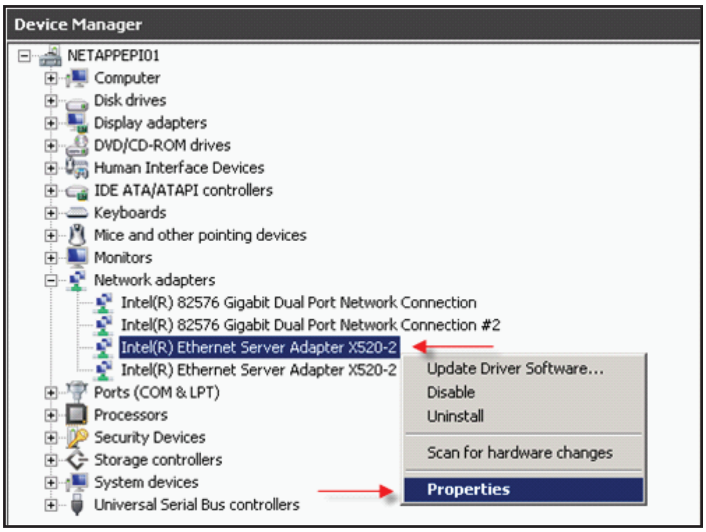


Figure 64

b. Verify driver.

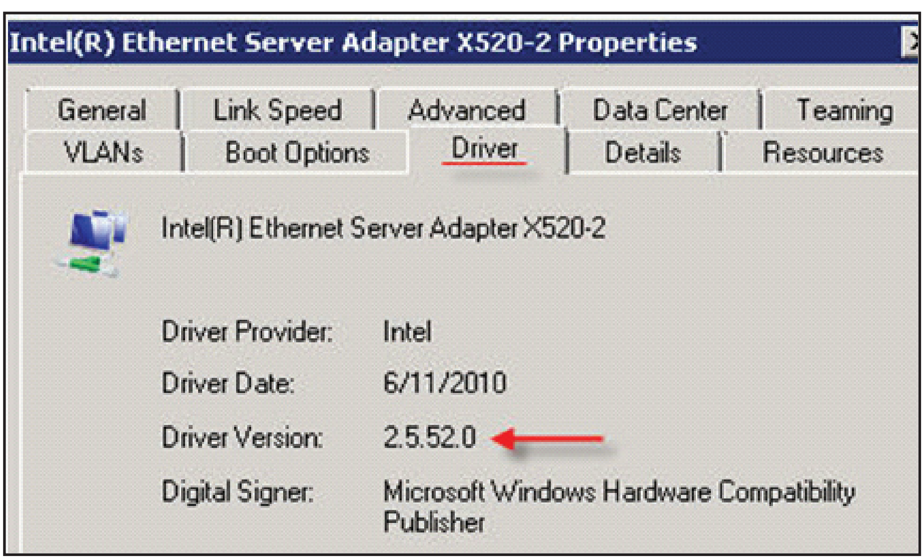


Figure 65

c. Verify that the Data Center tab is available. This verifies that the Data Center Bridging/Fibre Channel over Ethernet option is enabled. The FCoE Properties will navigate you to the Storage Controller Properties.

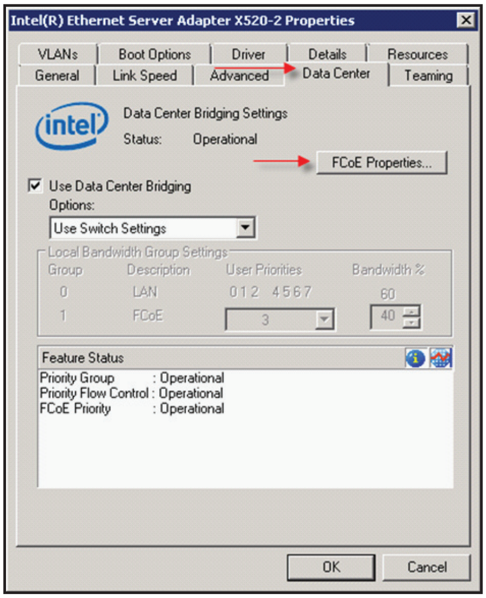


Figure 66

d. Through the Server Manager utility, browse to “Storage Controllers.” Right click and open “Properties.”

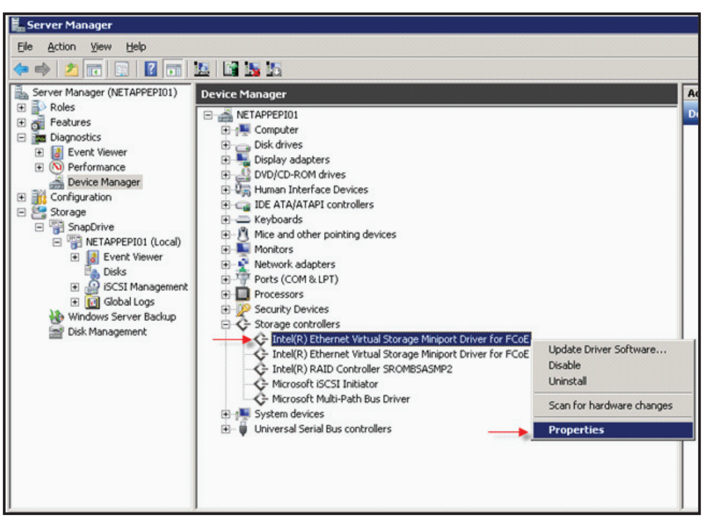


Figure 67

e. The Storage Controller properties are only available if the Data Center Bridging/Fibre Channel over Ethernet option is enabled. Most tabs are viewed to verify settings, but the primary tab we need is the "Fabric View" tab. It contains the Port WWN and Node WWN. These are required for configuration on the data center switch and the storage system.

***Please note that the secondary Hex locations are different (by design). This is done to reflect the VLAN ID for a particular connection.

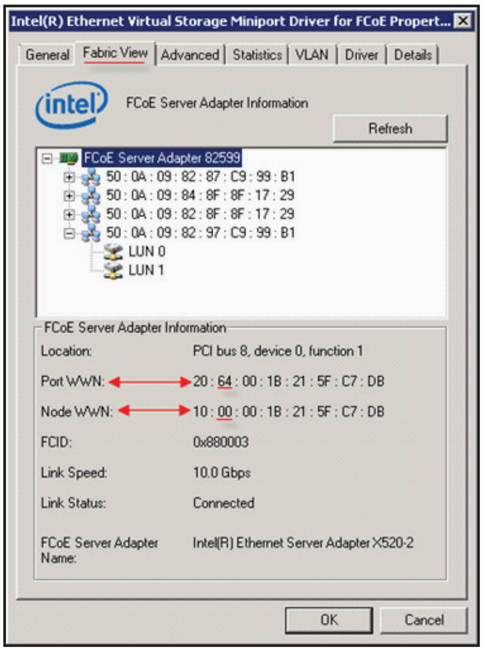


Figure 68

f. Other tabs to note:

i. Advanced: Port WWN. Port Properties button will navigate you to the Network Adapter Properties, and FC Port Storage Settings.

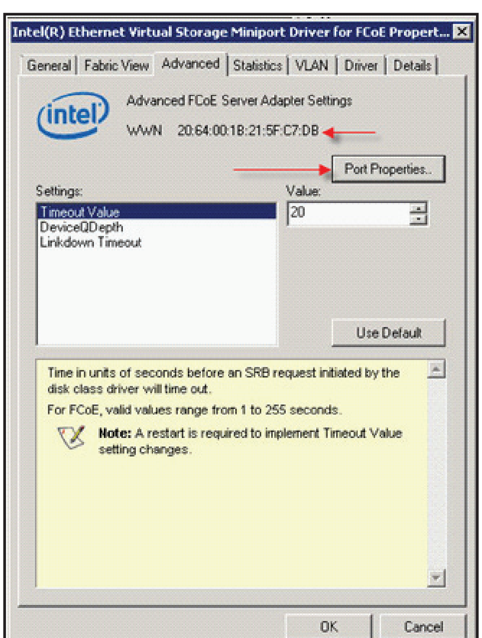


Figure 69

ii. Statistics: Give traffic information and reporting.

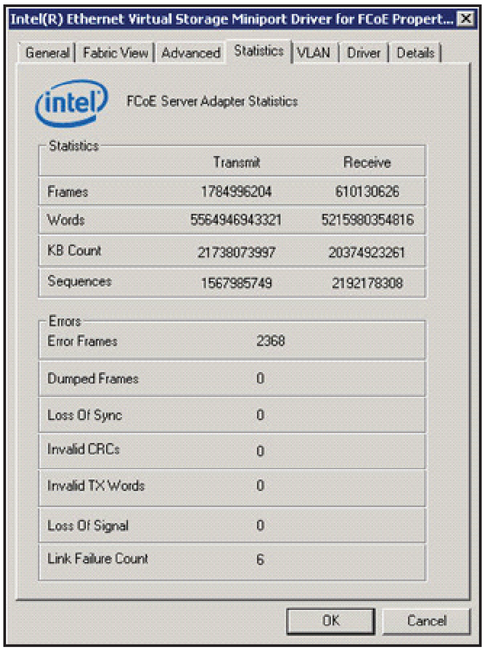


Figure 70

iii. VLAN: Give associated VLAN IDs.

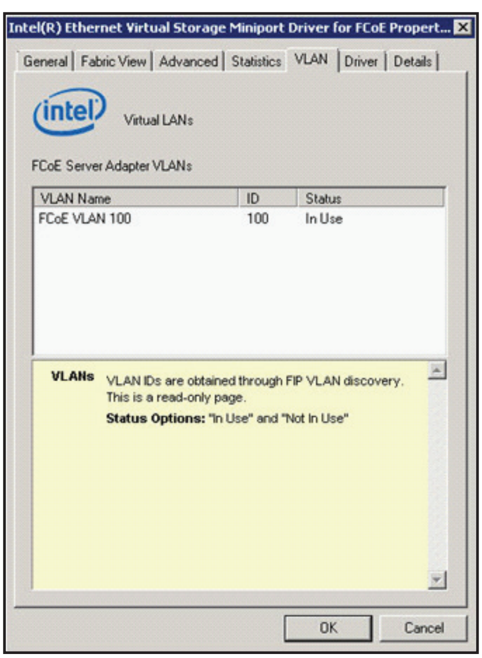


Figure 71

5. Stop! Proceed to the NetApp storage system to make the necessary changes to configure and assign a storage target to the server.
6. FCoE configuration on NetApp storage system:

   a. To create an FCoE LUN, open LUNs from the left menu of FilerView and select "Wizard" to open the LUN Wizard.

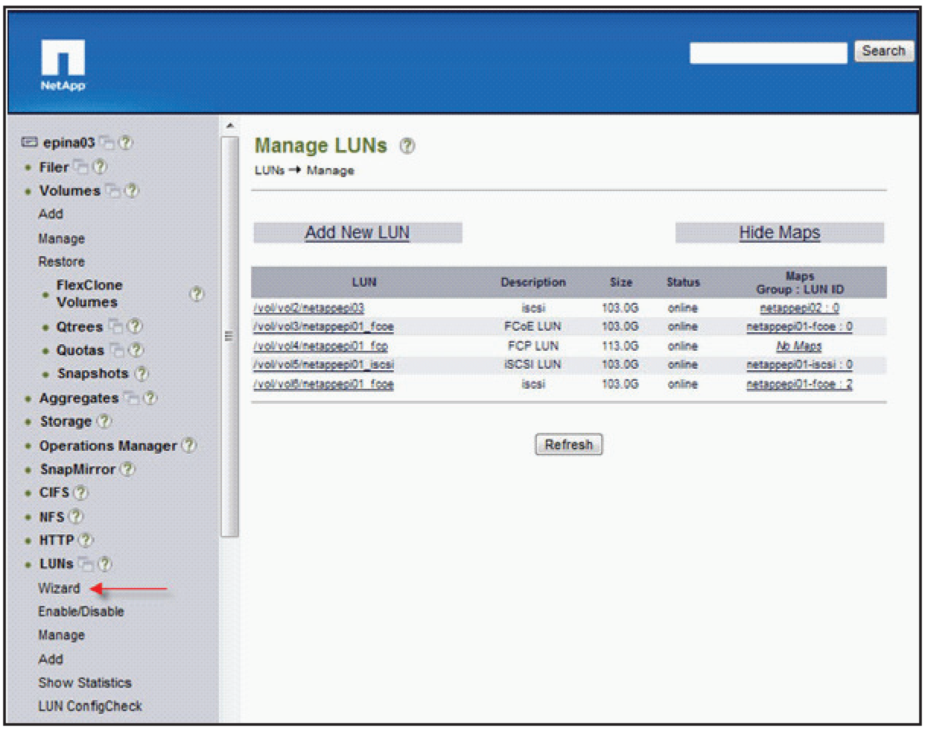


Figure 72

   b. Select "Next" at the LUN Wizard Welcome screen.

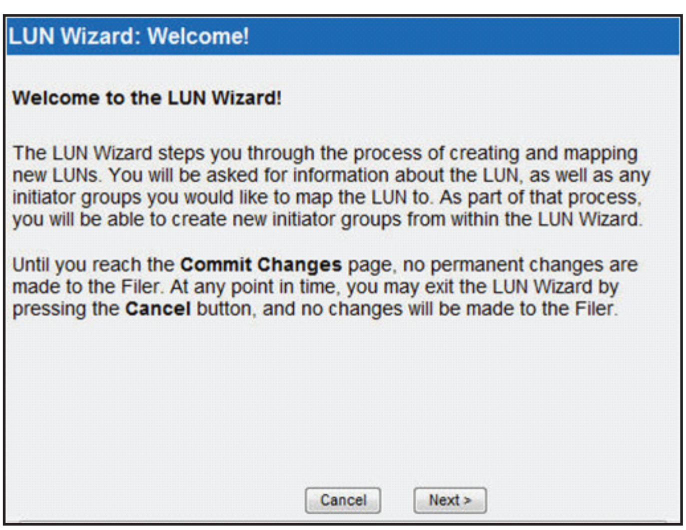


Figure 73

c. Set LUN configurations at the Specify LUN Parameters window. Path information is available through the Filer Manage Volumes window; the name is administrator defined. Set LUN size, protocol type, and select "Next."

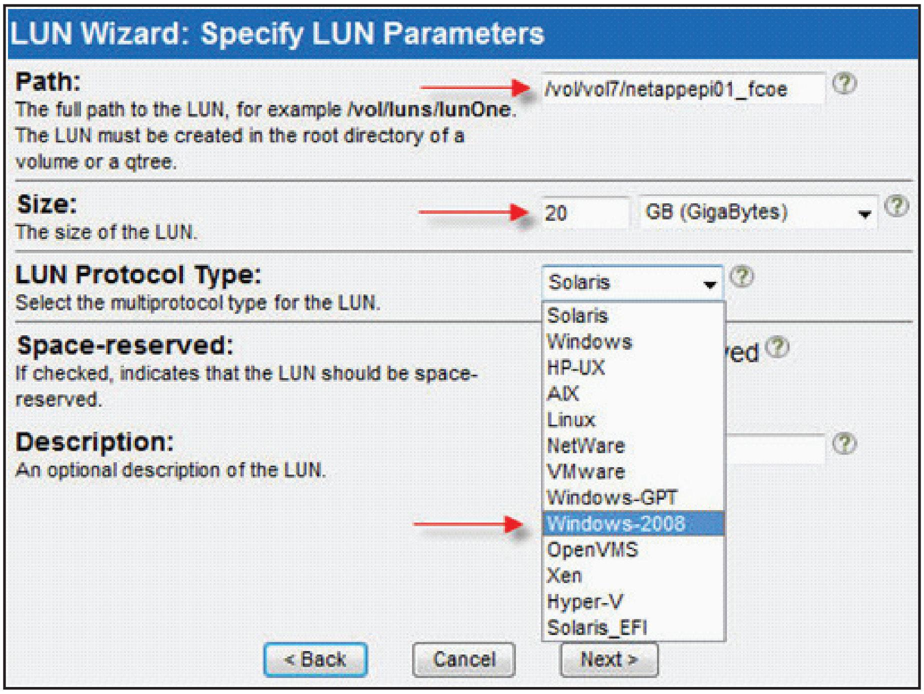


Figure 74

d. For a LUN to be activated, it must be added to an Initiator Group. Initiator Groups can be created during LUN creation or independently. This study used both methods; independently in FCoE and during LUN creation with iSCSI. Select "Next" to configure LUN without the Initiator Group, which will be created independently at a later point.

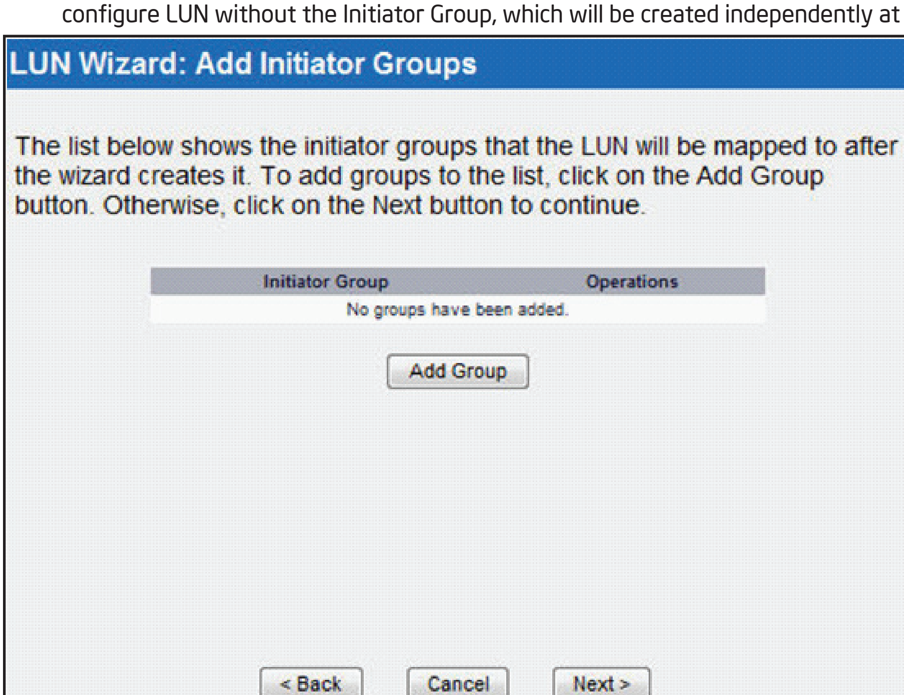


Figure 75

e. Select “Commit” to save the changes made in the LUN Wizard and close the success window.

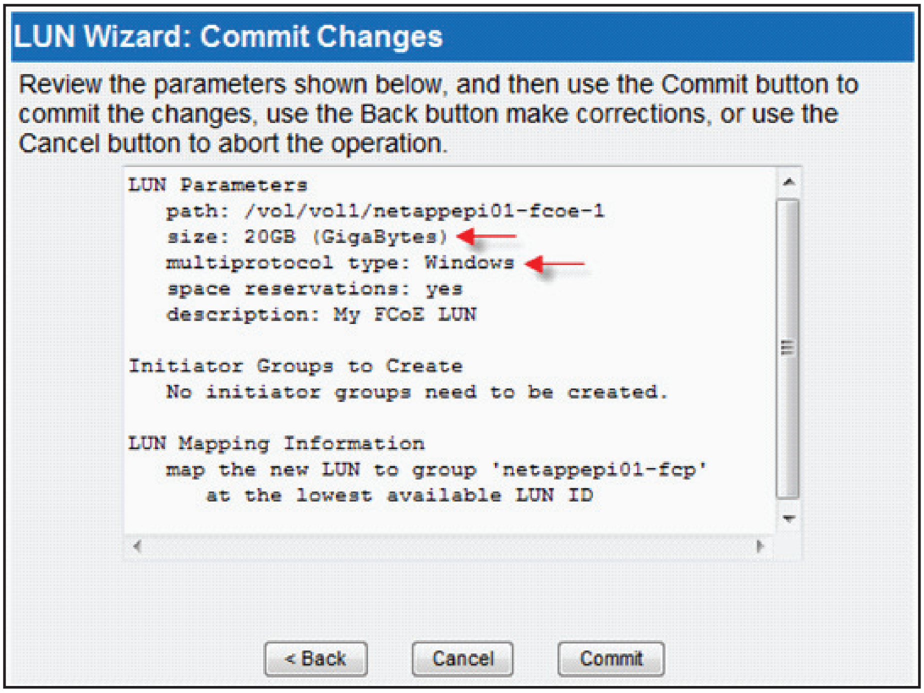


Figure 76

f. Verify the LUN’s creation in the Manage LUNs window. In this example, the LUN is created with no mapping to an Initiator Group, so it must be added for LUN usage. To add the mapping, select “No Maps” of the LUN that requires an Initiator Group.

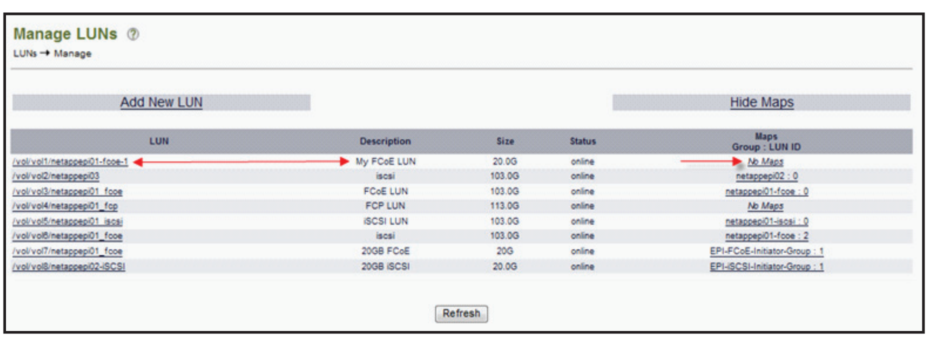


Figure 77

g. This will open the LUN Map window. Select “Add Groups to Map” for LUN, netappepi01-fcoe-1 and select “Apply.”

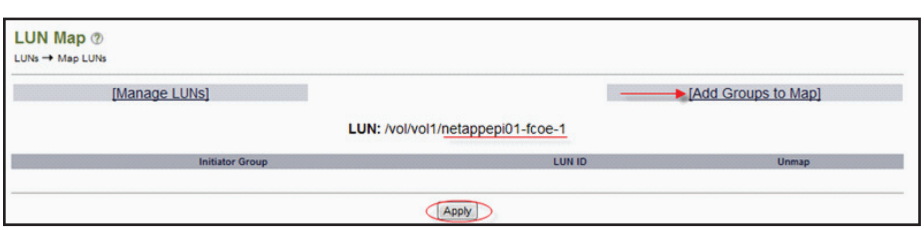


Figure 78

h. The LUN Map Add Groups window opens. Select the FCoE Initiator Group and select "Add."

Figure 79

i. Create a LUN ID number or if left blank, the system will create it, and select "Apply."

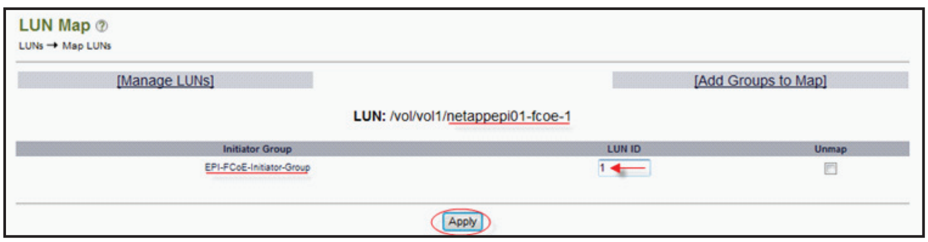


Figure 80

j. Verify the LUN has been mapped to the correct group in the Manage LUN window.

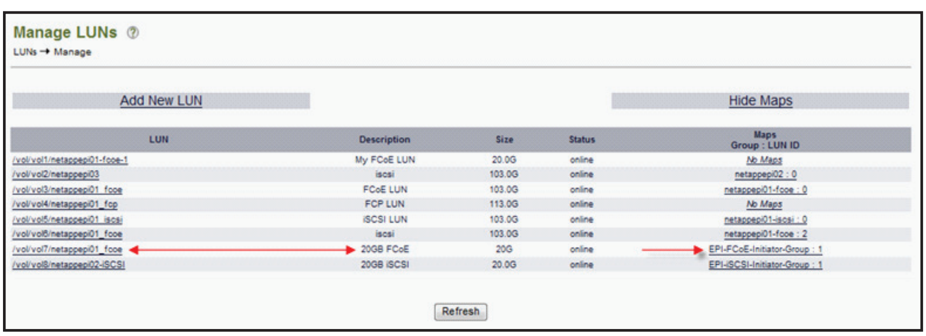


Figure 81

7. **Stop!** The LUN is now created, so move back to the server and add the imported drive.
8. Initializing Drives: If network and storage parameters are correctly configured, LUNs will be applied to the server. These drives will act like hot-swapped, imported drives which need to be brought online, initialized, and formatted for use.

   a. Right click the unknown disk and bring it "Online."

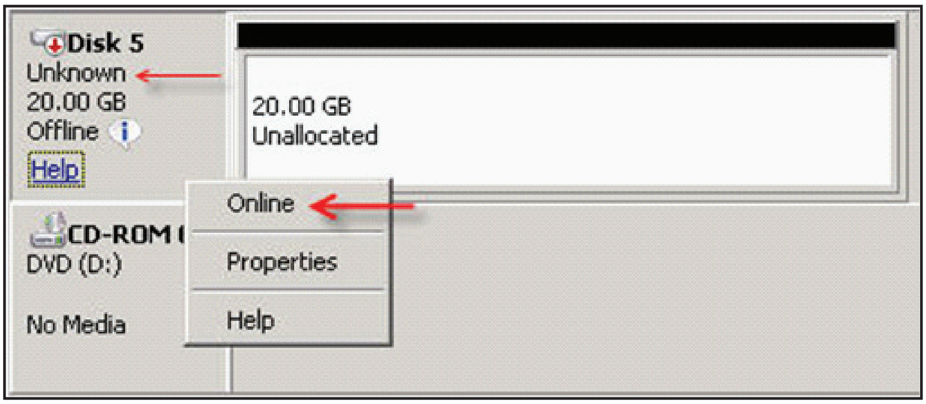


Figure 82

   b. Right click the unknown disk to initialize the disk.

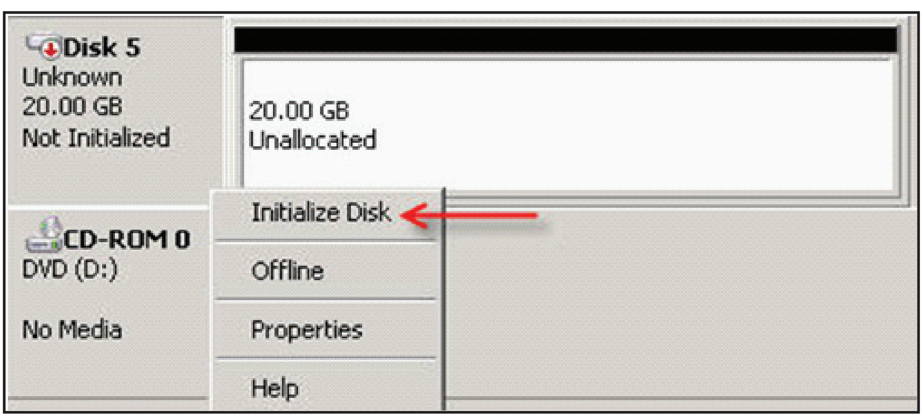


Figure 83

   c. Create a new simple volume and format for use.

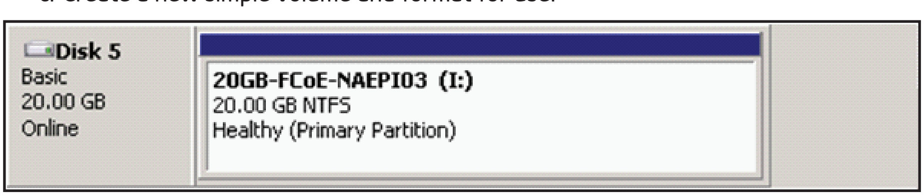


Figure 84

Iometer Test Tool Configuration Details

1. On your windows server host, download and install Iometer. You can download Iometer at http://www.iometer.org.
2. Before you launch Iometer, make sure your iSCSI/FCoE LUNS are mapped to the Windows host.
3. From server's start menu, navigate to "Iometer."

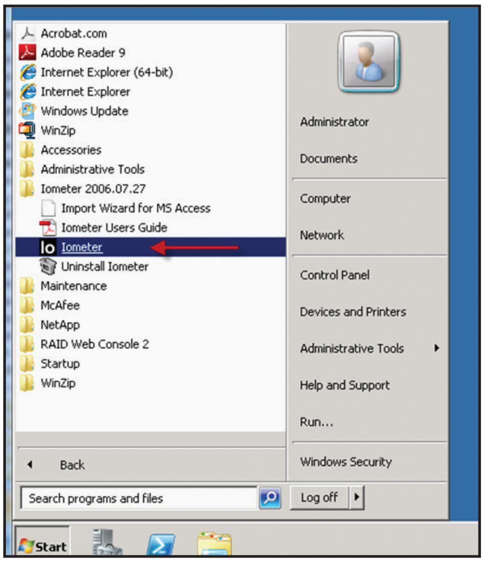


Figure 85

4. The default screen will appear as shown in Figure 86.

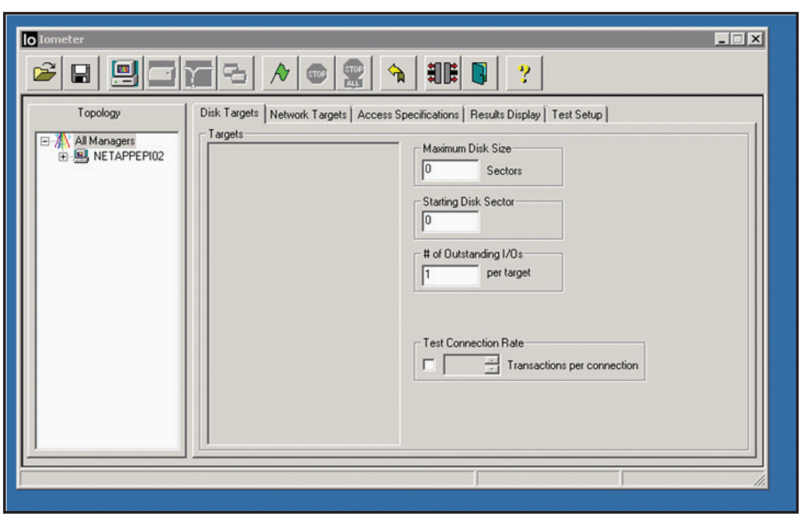


Figure 86

5. Now go to the "Topology" section and select your Windows server host name. Expand the manager and you will see a list of workers. By default, the number of workers is equal to the number of processor cores you have on the server (24 in this example).

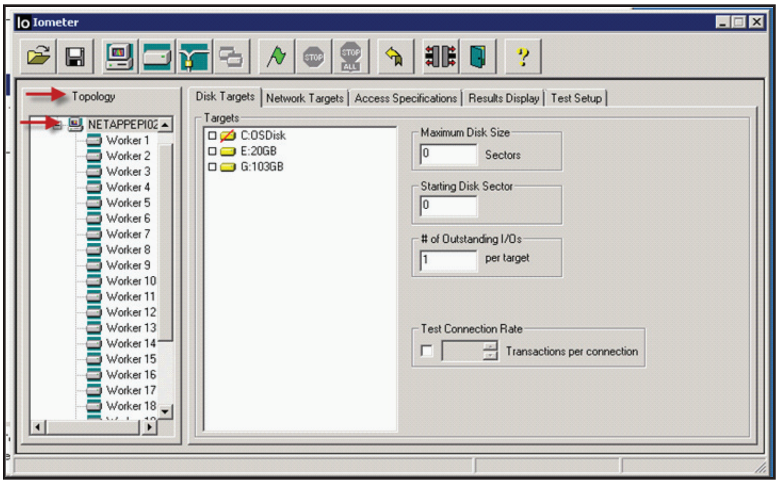


Figure 87

6. Select the manager: in our example, "NETAPPEPI02" and the "Disk Targets" tab should be selected by default (if not, select it). By default the values in the fields "Maximum Disk Size" and "# of Outstanding I/Os" are set to 0 and 1 respectively. Change these values to those shown below:

   Maximum Disk Size = 10000 sectors

   # of Outstanding I/Os = 8 per target.

   Note 1- Iometer creates a file called "iobw.tst" in the iSCSI/FCoE LUN based on the value you enter in the "Maximum Disk Size" (and runs the tests with this file size). If you use the default "0," Iometer will create a file equal to the size of your LUN. In this example, since we used a value of 10000, a file of 5 MB was created.

   Note 2 – In our testing, we tried various values for "# of Outstanding I/Os" and we found that "8" gave the optimal performance. This number is dependent on the queue depth of your target.

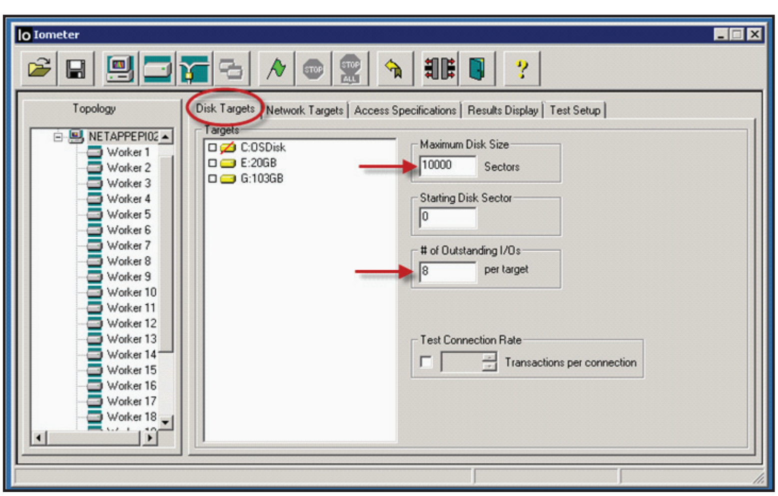


Figure 88

7. In the "Disk Targets" section, select the disk target you want to run the tool against. In order to use more than 1 worker (which is default), you need to select the disk target for each worker manually, as shown below.

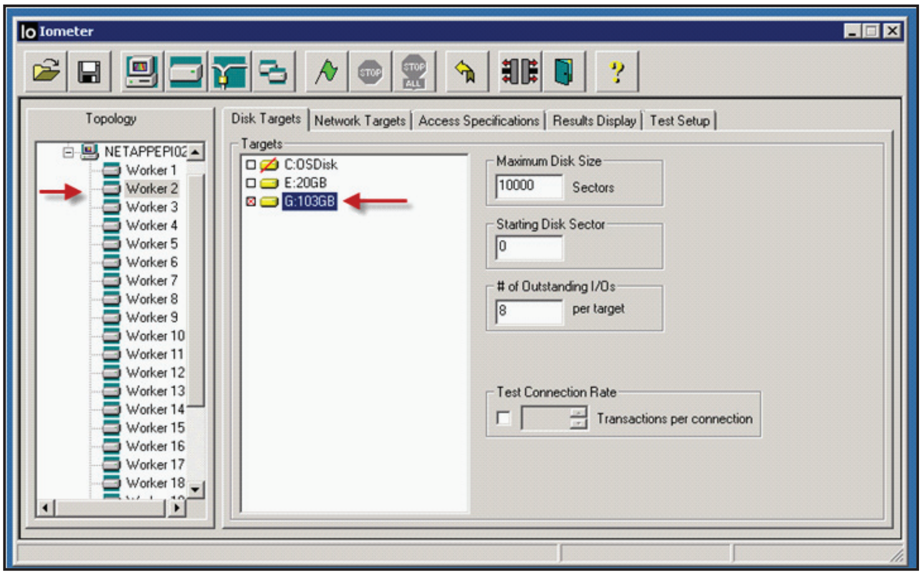


Figure 89

8. Select the "Access Specifications" tab in the Iometer window and select "New" to create a new test.

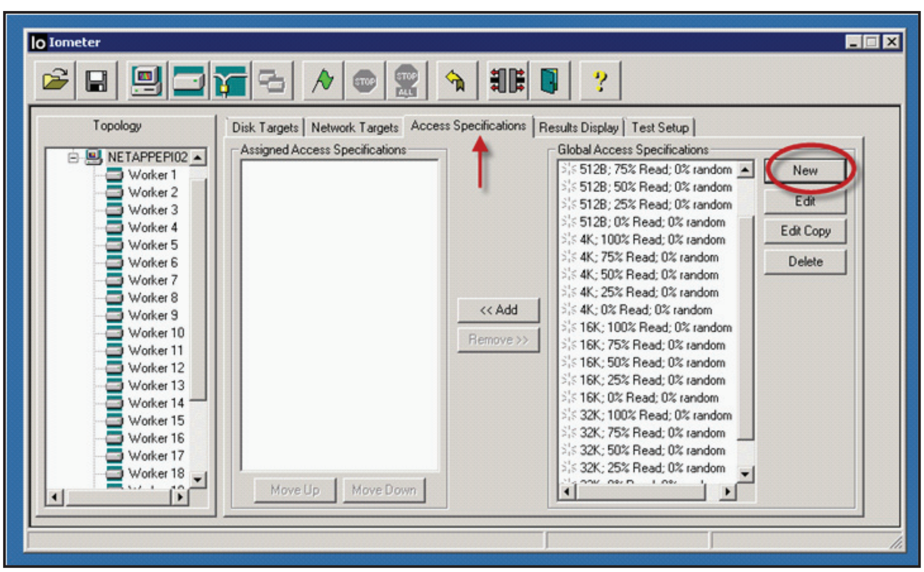


Figure 90

9. In the new window, specify the test you want to run. Below is a sample for 4 K 100% sequential read.

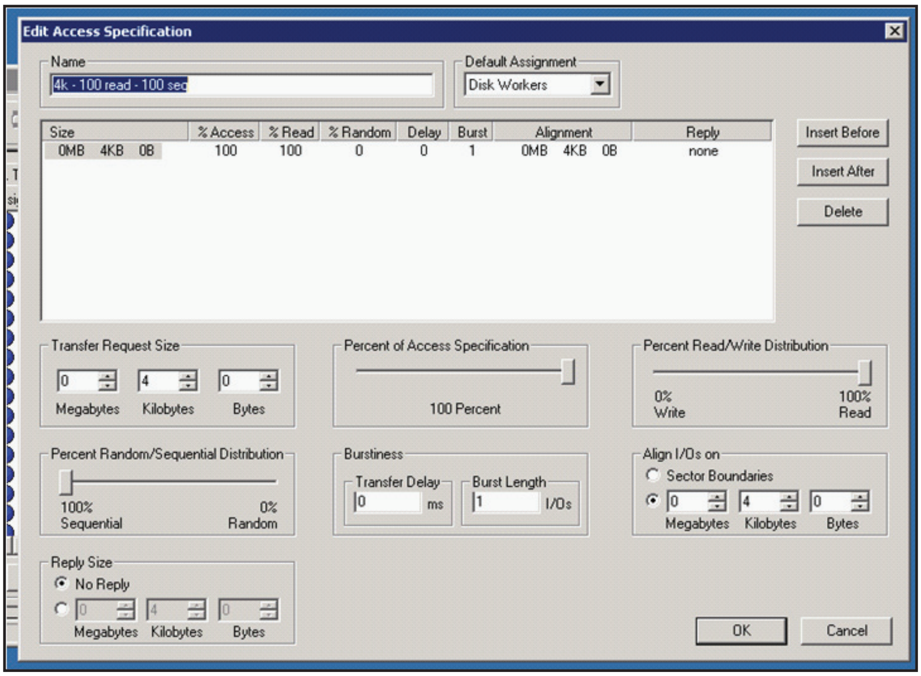


Figure 91

10. In our testing we used varying request sizes (4 KB to 1024 KB), and for all these request sizes we used the following combinations:

- 100% read – 100% sequential
- 100% write – 100% sequential
- 100% read – 100% random
- 100% write – 100% random
- 50% read – 100% sequential
- 50% read – 100% random
- 50% read – 50% random

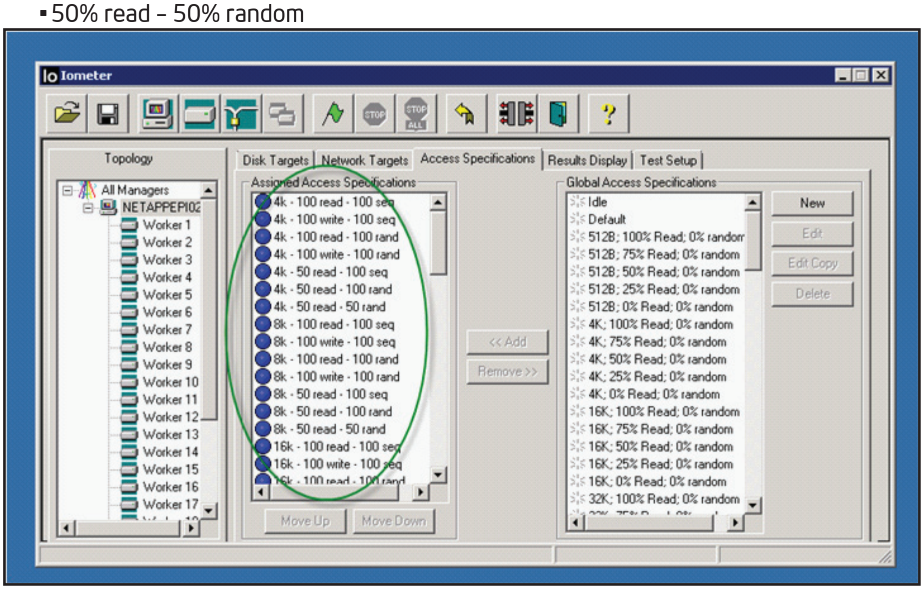


Figure 92

11. On the Iometer window, select the test setup section so that you can specify the test duration, ramp up time, etc. Figure 93 shows the sample.

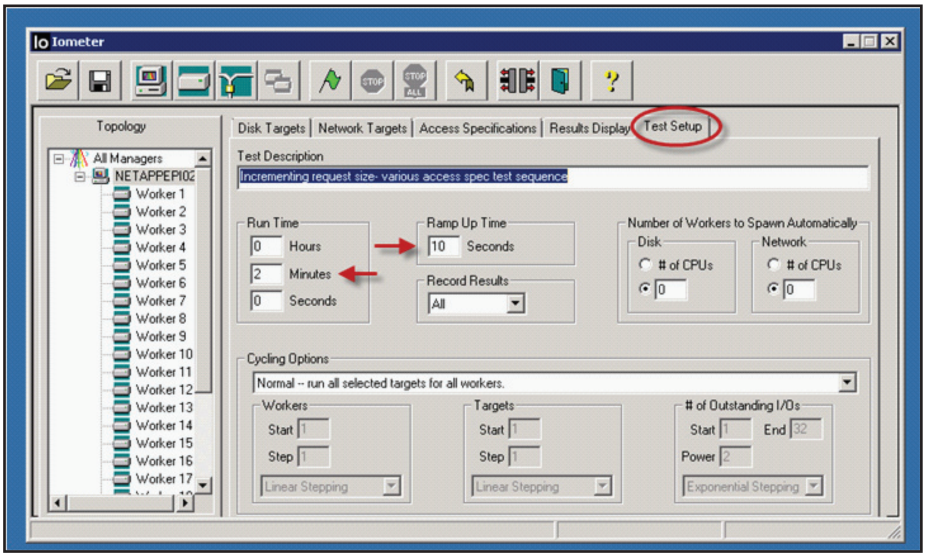


Figure 93

12. After the setup, start the test by clicking on the green flag.

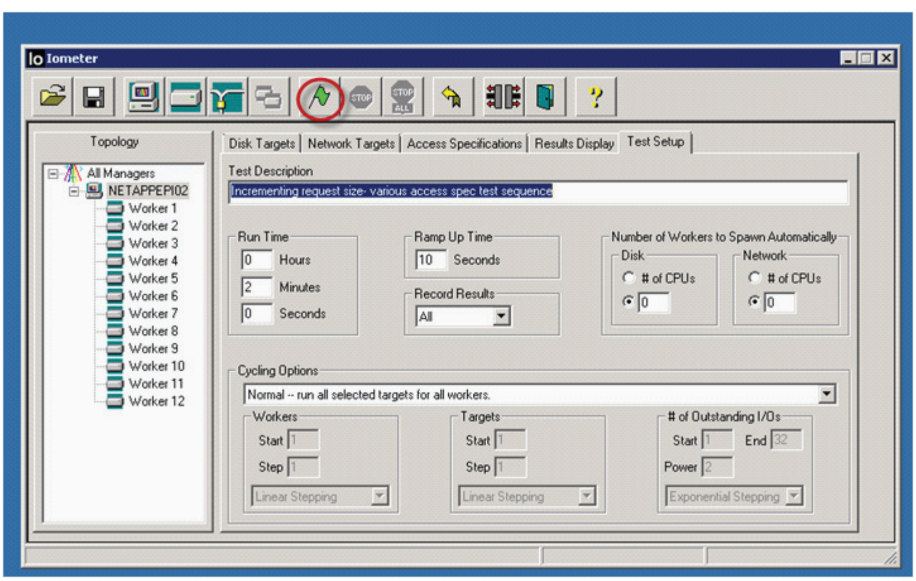


Figure 94

13. Figure 95 shows the results screen.

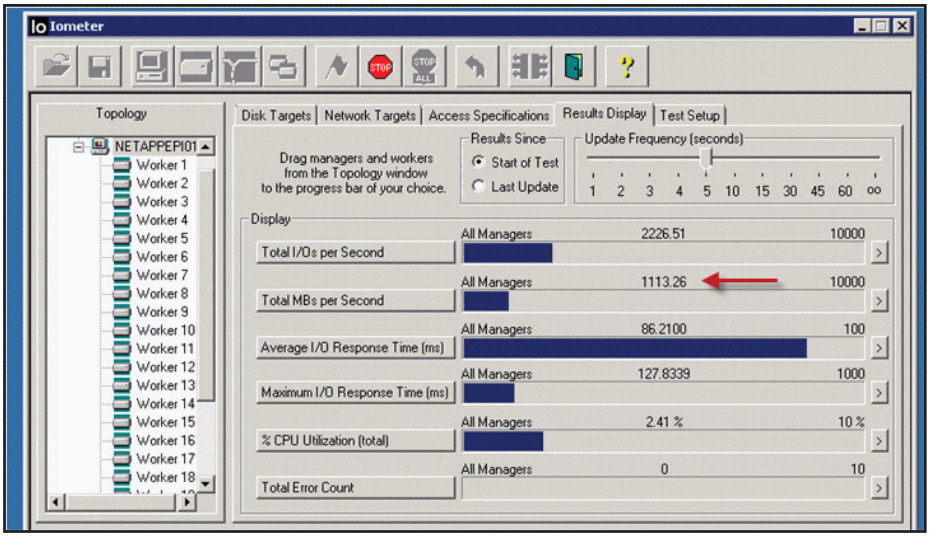


Figure 95

iSCSI Functionality and Performance Results

The iSCSI native OS initiator running over a 10 GbE end-to-end solution was fully functional and performed very well.

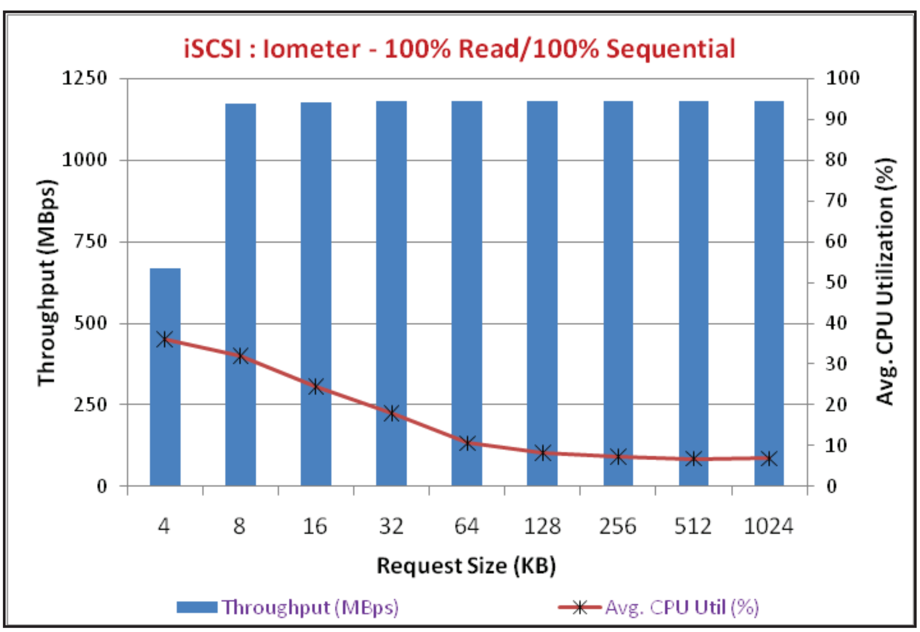


Figure 96

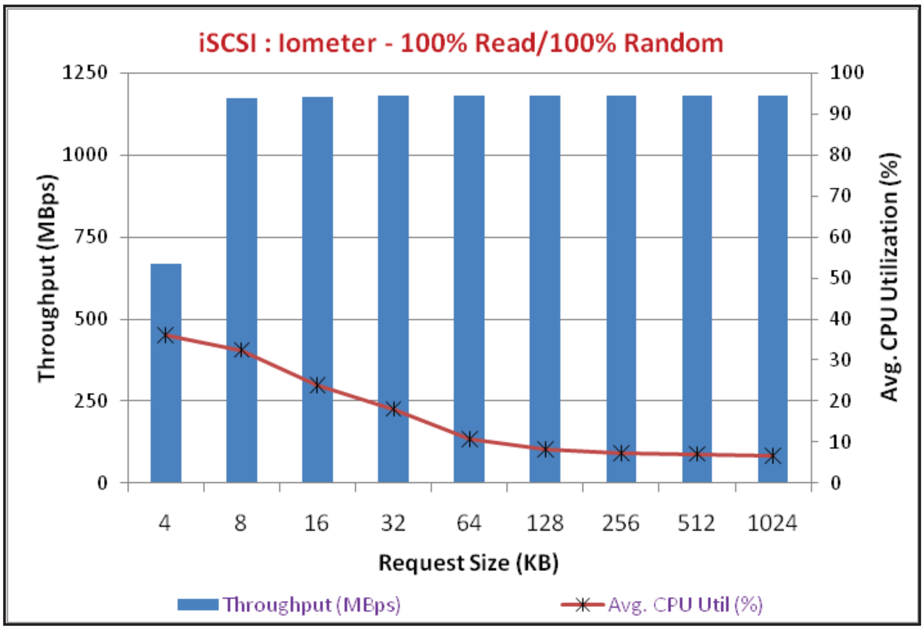


Figure 97

Throughput approached the 10 Gbps line rate for the 100% read tests at block sizes of 8 KB and above. Processor utilization peaked ~37% for the 4 KB block size, and declined to ~7% for the larger block sizes. No difference was observed for random vs. sequential access patterns. This is attributable to the large Flash Cache in the NetApp system.

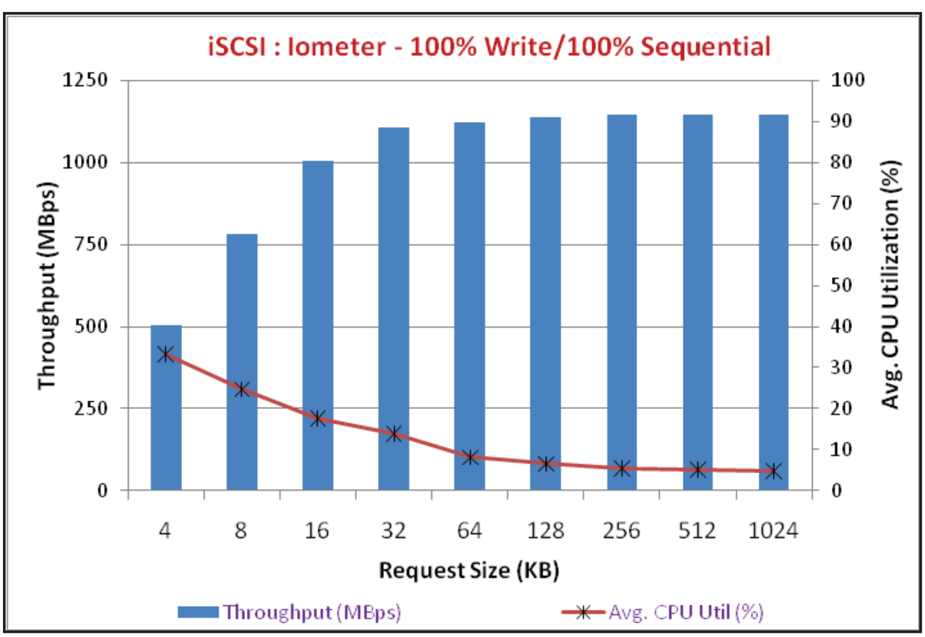


Figure 98

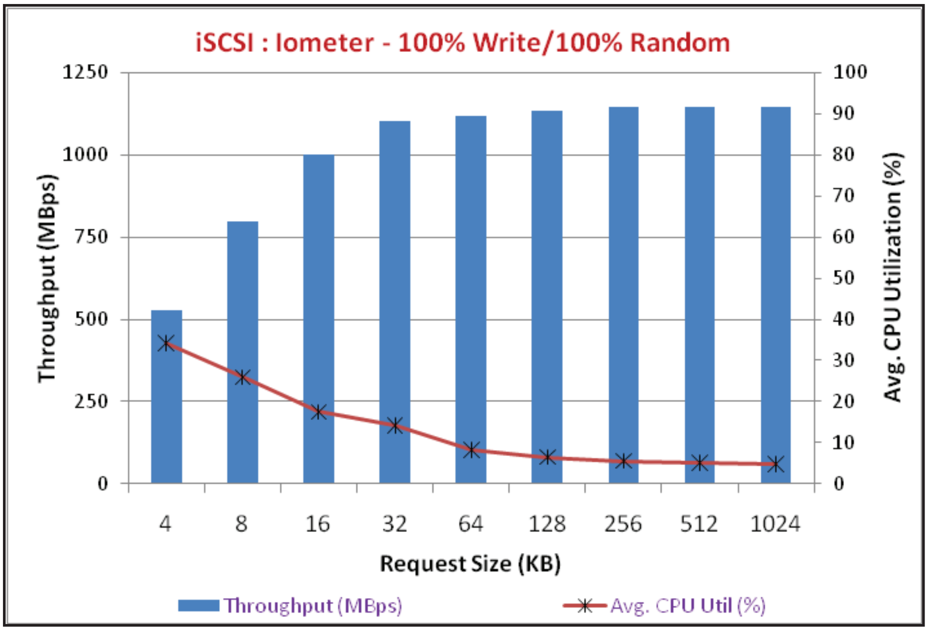


Figure 99

Throughput approached line rate for the 100% write tests at block sizes of 32 KB and above. Processor utilization peaked ~33% for the 4 KB block size, and declined to ~5% for the larger block sizes. No difference was observed for random vs. sequential access patterns. This is attributable to the large Flash Cache in the NetApp system.

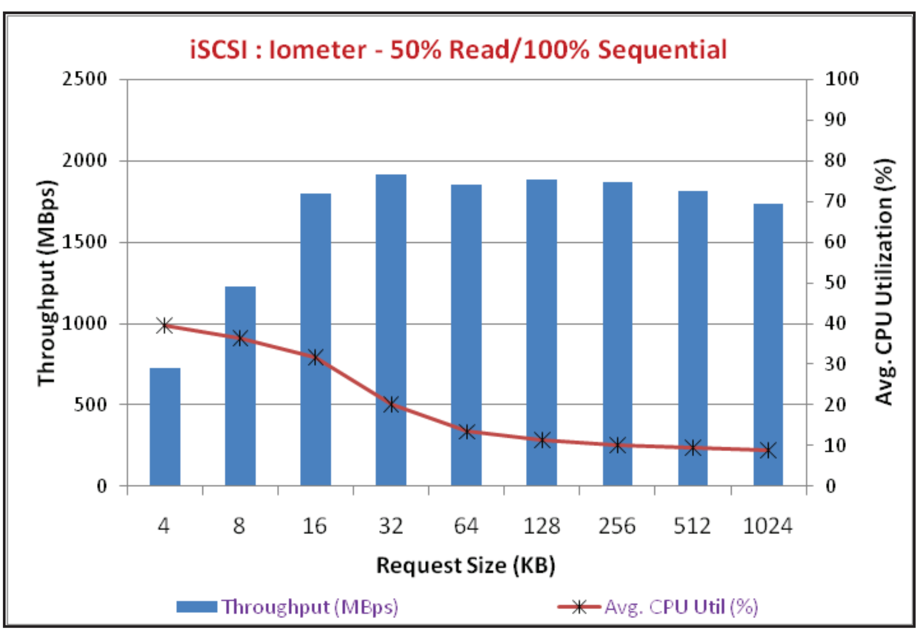


Figure 100

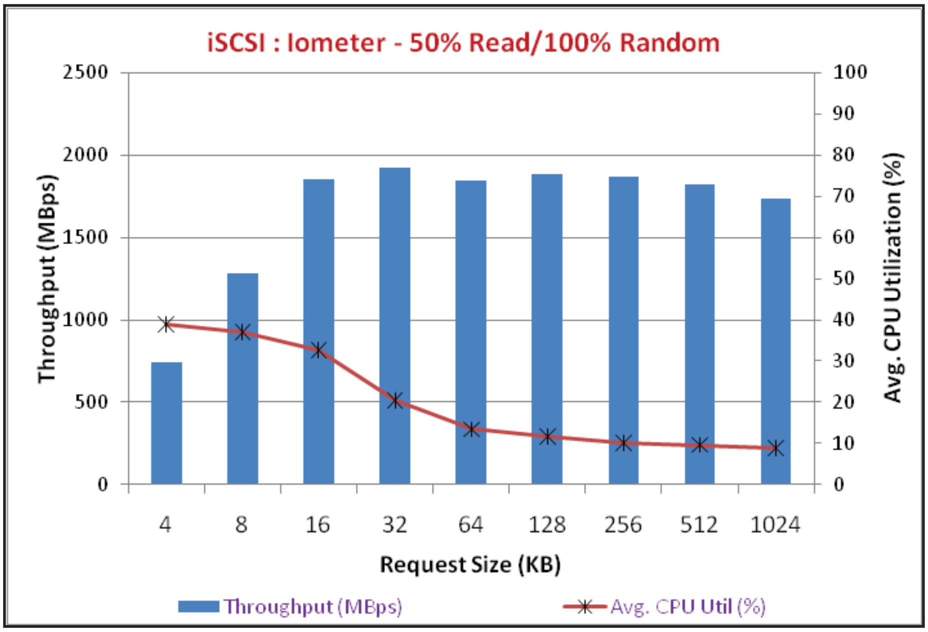


Figure 101

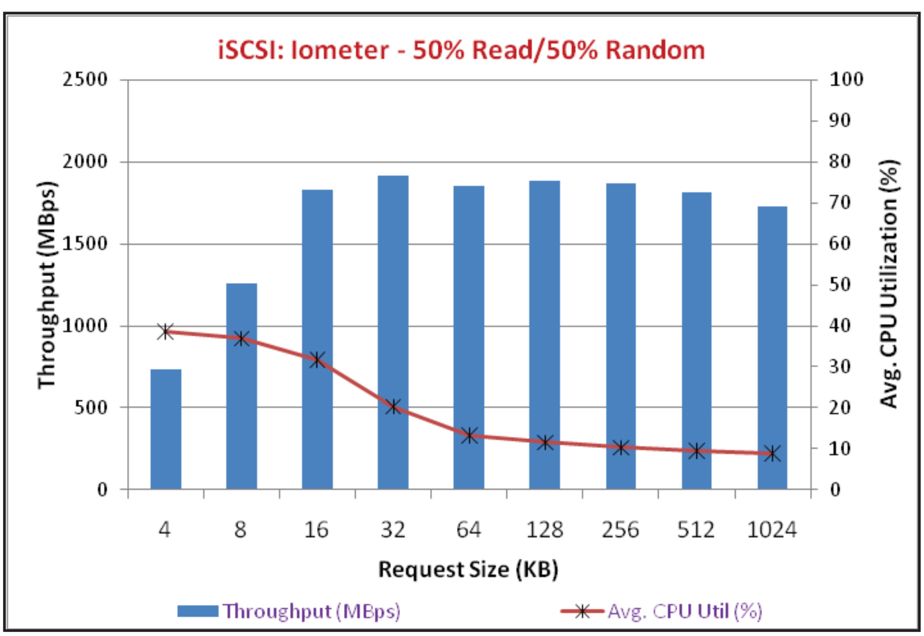


Figure 102

Throughput peaked at ~1900 MB/sec in the read/write 50% tests even though read and write individually both peaked at over 1100 MB/sec. We are not sure why read/write is below the sum of read and write individually. We will have to do some more performance tuning to optimize this. Processor utilization peaked ~39% for the 4 KB block size and declined to ~9% for the larger block sizes. No difference was observed for random vs. sequential access patterns. This is attributable to the large Flash Cache in the NetApp system.

FCoE Functionality & Performance Results

The FCoE native OS initiator running over a 10 GbE end-to-end solution was fully functional and performed very well.

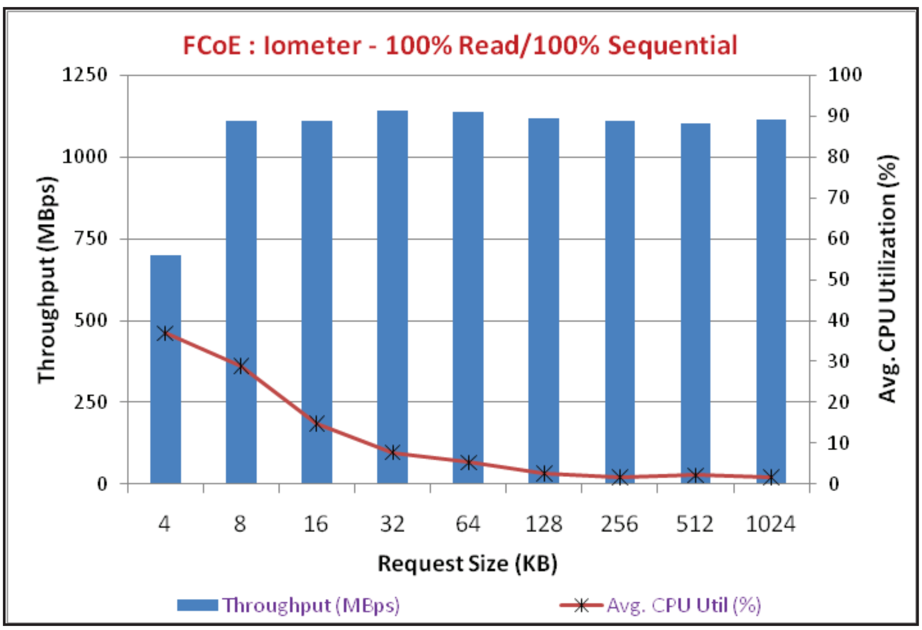


Figure 103

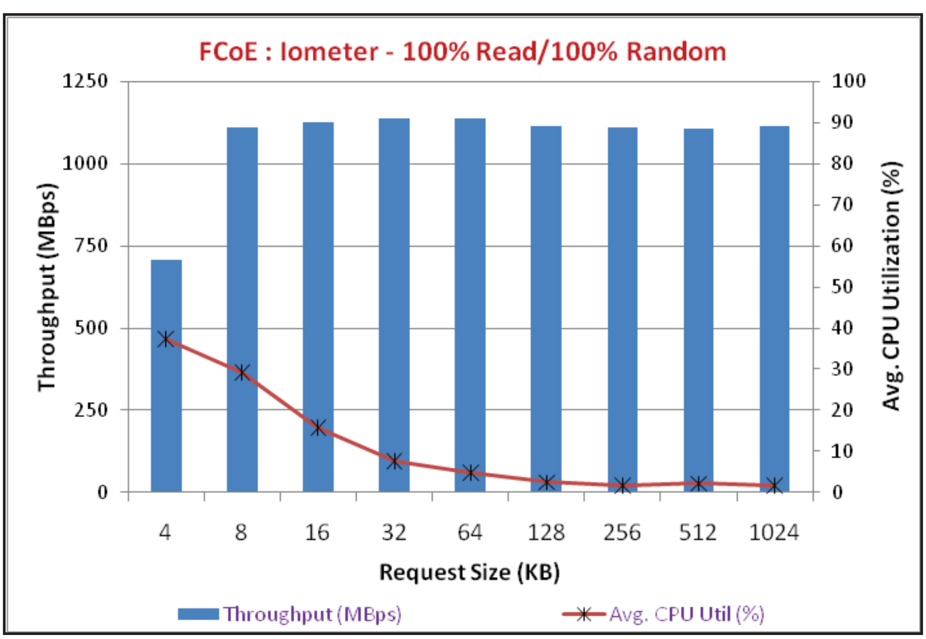


Figure 104

Throughput approached the 10 Gbps line rate for the 100% read tests at block sizes of 8 KB and above. Processor utilization peaked ~37% for the 4 KB block size, and declined to ~2% for the larger block sizes. No difference was observed for random vs. sequential access patterns. This is attributable to the large Flash Cache in the NetApp system.

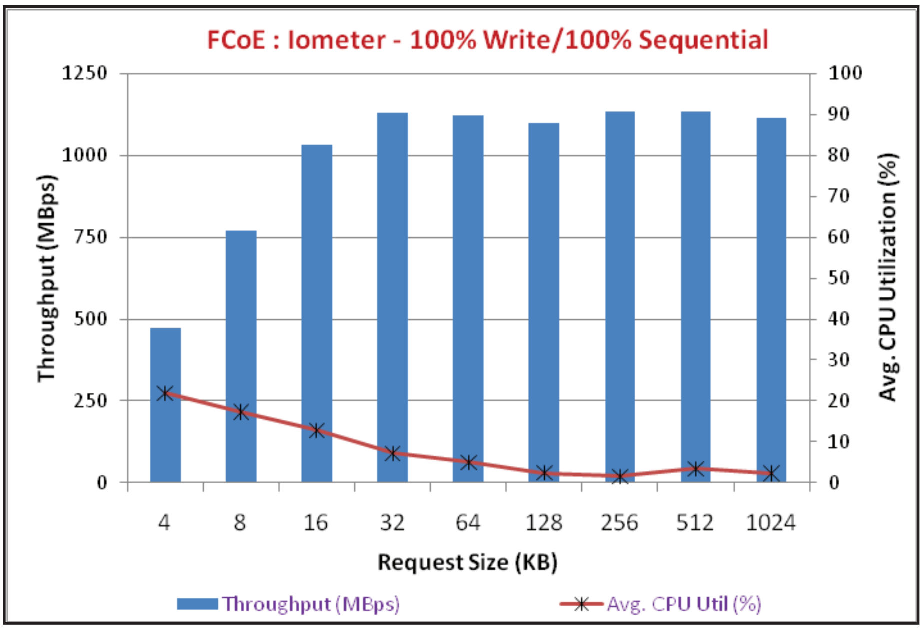


Figure 105

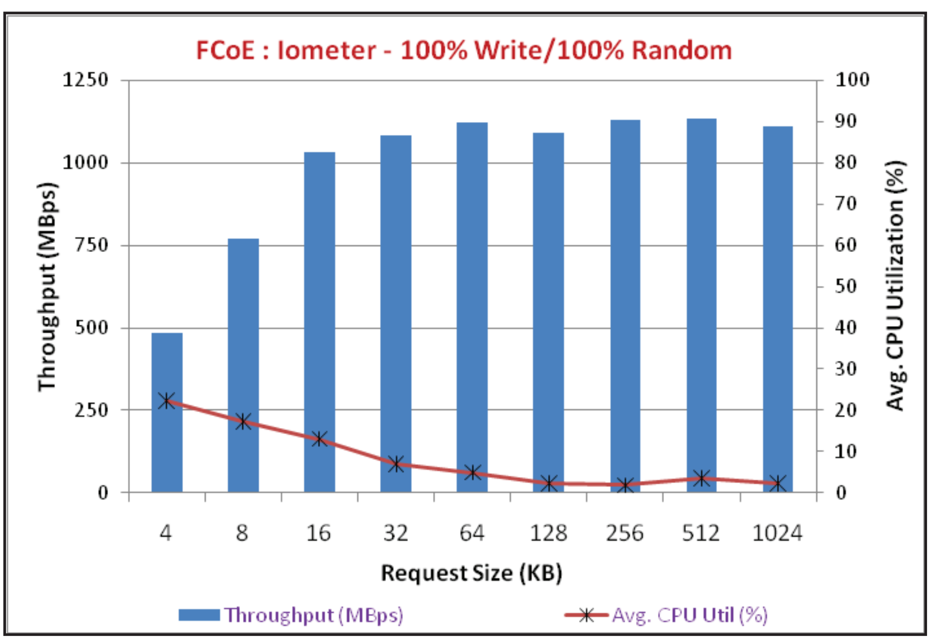


Figure 106

Throughput approached line rate for the 100% write tests at block sizes of 32 KB and above. Processor utilization peaked ~22% for the 4 KB block size, and declined to ~2% for the larger block sizes. No difference was observed for random vs. sequential access patterns. This is attributable to the large Flash Cache in the NetApp system.

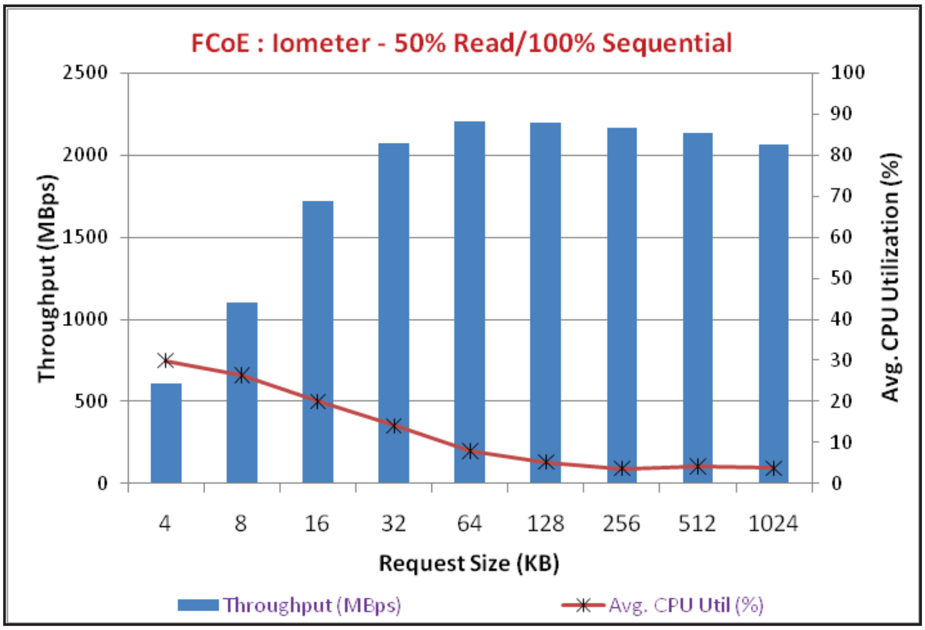


Figure 107

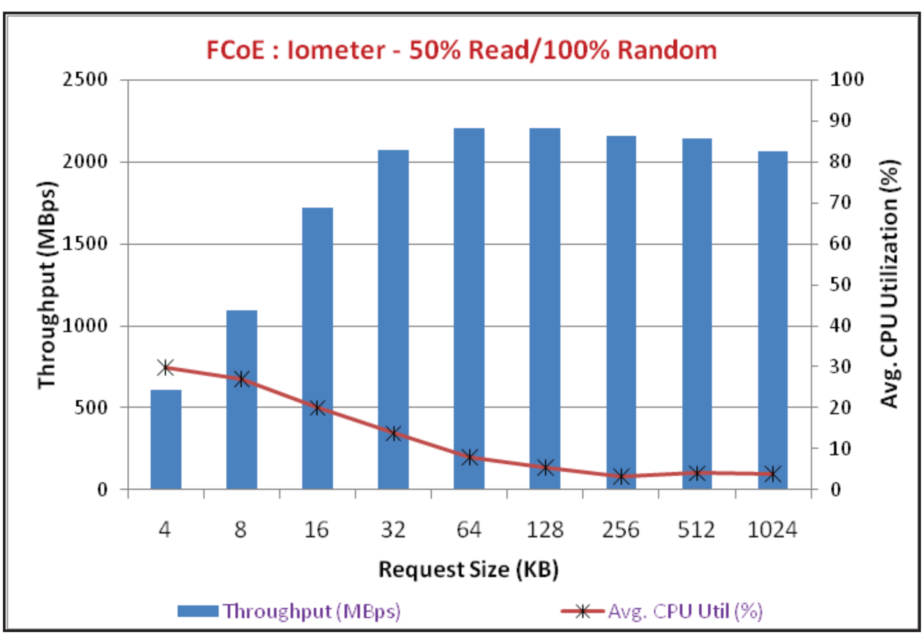


Figure 108

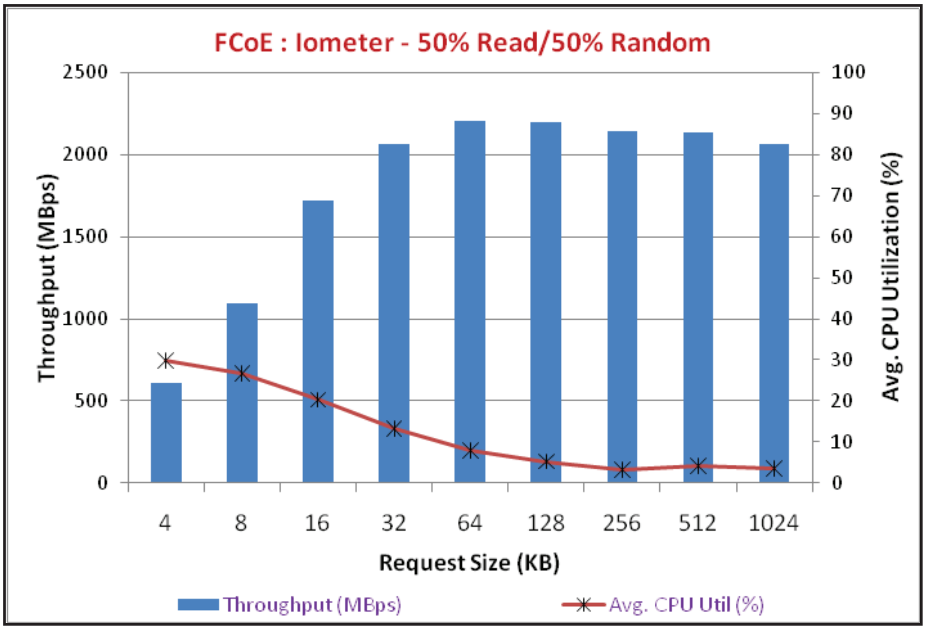


Figure 109

Throughput peaked at ~2200 MB/sec in the read/write 50% tests. This result is nearly the sum of our individual read and write test results, which is better than the iSCSI test. Processor utilization peaked ~30% for the 4 KB block size, and declined to ~4% for the larger block sizes. No difference was observed for random vs. sequential access patterns. This is attributable to the large Flash Cache in the NetApp system.

Analysis

Both the iSCSI and FCoE native OS initiators running over a 10 GbE end-to-end solution were fully functional and performed very well.

Throughput approached the 10 Gbps line rate for the 100% read tests at block sizes of 8 KB and above.

Throughput approached line rate for the 100% write tests at block sizes of 32 KB and above.

Throughput peaked at ~2200 MB/sec in the read/write 50% tests for FCoE, but only ~1900 MB/sec for iSCSI. This may be attributable to the TCP/IP stack overhead and possibly reflects an imperfectly tuned system. We continue to work on improvement of this result now.

Processor utilization on the server was somewhat higher for the iSCSI cases. This may be attributable to the TCP/IP stack overhead and possibly reflects an imperfectly tuned system.

No difference was observed for random vs. sequential access patterns. This is attributable to the large Flash Cache in the NetApp system.

One key factor in the success of native OS initiator solutions for unified networking is the substantially increased processing performance of the Intel® Xeon® processor 5500 & 5600 series product lines available for the last few years. Another key factor is the performance optimization in the latest Intel X520 series of NICs. These devices provide hardware support to spread the network processing load across the many cores of modern server platforms. They also have protocol specific hardware optimizations for both iSCSI and FCoE.

Because the test bed was available to us for only a few weeks, we could not do all of the tests and performance tuning that we would have liked to do. However, we have demonstrated near 10 Gbps line speed performance for both iSCSI and FCoE.

Next Steps

Following this research, next steps include:

- Optimization of tuning in iSCSI cases to maximize performance
- Configuration of NFS & CIFS protocols: testing & comparison with iSCSI & FCoE results
- Extension of testing to other OS & VMM environments beyond Windows 2008
- Extension of testing to multipath topologies to find the maximum performance possible

Conclusion

The simplicity, flexibility, and familiarity of the 10 Gigabit Ethernet make it the ideal fabric for a unified data center network fabric. It enables new compute deployment models, including that of cloud computing, which will deliver more intelligent, responsive data centers, and greater business agility. Through their product and industry leadership, Intel and NetApp help IT organizations transition to 10 GbE and build next-generation data center infrastructures.

Users have a choice of several viable options for unified networking protocols running on 10 GbE. All of these choices run very well with native OS initiators running on standard 10 GbE NICs. The use of native OS initiator solutions and standard NICs for unified networking delivers cost-effective solutions, ubiquitous deployment, and flexible choices of protocols.

We encourage the reader to explore 10 GbE unified networking in their own labs to determine the best choices for their own applications and environments.

Additional Information

1. Intel® Ethernet X520 Server Adapters: http://www.intel.com/Products/Server/Adapters/X520/ethernet-X520-overview.htm
2. Intel® Xeon® Processor Family: http://www.intel.com/itcenter/products/xeon/index.htm
3. Intel® Server System SR2600UR: http://www.intel.com/products/server/systems/sr2600ur/sr2600ur-overview.htm
4. NetApp SnapDrive* admin guide: http://now.netapp.com/knowledge/docs/snapdrive/relsnap63/pdfs/admin.pdf
5. NetApp RAID-DP*: http://www.netapp.com/us/products/platform-os/raid-dp.html ; http://www.netapp.com/us/library/technical-reports/tr-3298.html
6. NetApp Flash Cache*: http://www.netapp.com/us/products/storage-systems/flash-cache/
7. Data Center Bridging: http://www.ieee802.org/1/pages/dcbridges.html
8. Fibre Channel over Ethernet: http://en.wikipedia.org/wiki/FCoE
9. Open FCoE: http://open-fcoe.org/
10. Internet Small Computer System Interface: http://en.wikipedia.org/wiki/Iscsi

Appendix A: NetApp Command Line Setup

For CLI command reference see http://now.netapp.com/NOW/knowledge/docs/ontap/rel80/pdfs/ontap/cmdref1.pdf and http://now.netapp.com/NOW/knowledge/docs/ontap/rel80/pdfs/ontap/cmdref2.pdf. Below are the Data ONTAP command-line equivalents to the FilerView-based setup dialogues found in the "NetApp Environment Setup" section of this document.

iSCSI interface management:

```
> iscsi interface disable e2b
> iscsi interface enable e0c
```

Aggregate creation:

```
> aggr create -t EPIaggr2 raid_dp 3
```

Volume creation:

```
> vol create EPIVol01 EPI-Aggr2 42M
```

Igroup creation and population (FCoE)

```
> igroup create -f -t windows EPI-FCoE-Initiator-Group
> igroup add EPI-FCoE-Initiator-Group 20:64:00:1b:21:56f:c7:db
```

Igroup creation and population (iSCSI)

```
> igroup create -I -t windows EPI-iSCSI-Initiator-Group
> igroup add EPI-iSCSI-Initiator-Group iqn.1991-05.com.microsoft:netappepi01
```

LUN creation and mapping (FCoE)

```
> lun create –s 20G –t windows-2008 –o noreserve /vol/vol7/netappepi01_fcoe
> lun map /vol/vol7/netappepi01_fcoe EPI-FCoE-Initiator-Group 0
```

LUN creation and mapping (iSCSI)

```
> lun create –s 20G –t windows-2008 –o noreserve /vol/vol7/netappepi02-iSCSI
> lun map /vol/vol7/netappepi01-iSCSI EPI-iSCSI-Initiator-Group 0
```

Appendix B: NetApp Jumbo Frame Confirmation through GUI

1. Select "Manage Interfaces" under Network from the left menu of FilerView > "Select Modify" to view or change parameters of the interface.

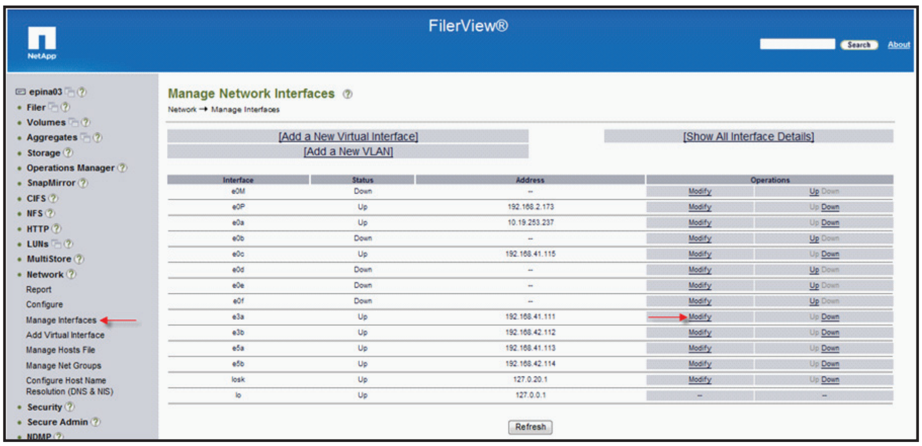


Figure 110

2. Interface e3a shows that jumbo frames are on.

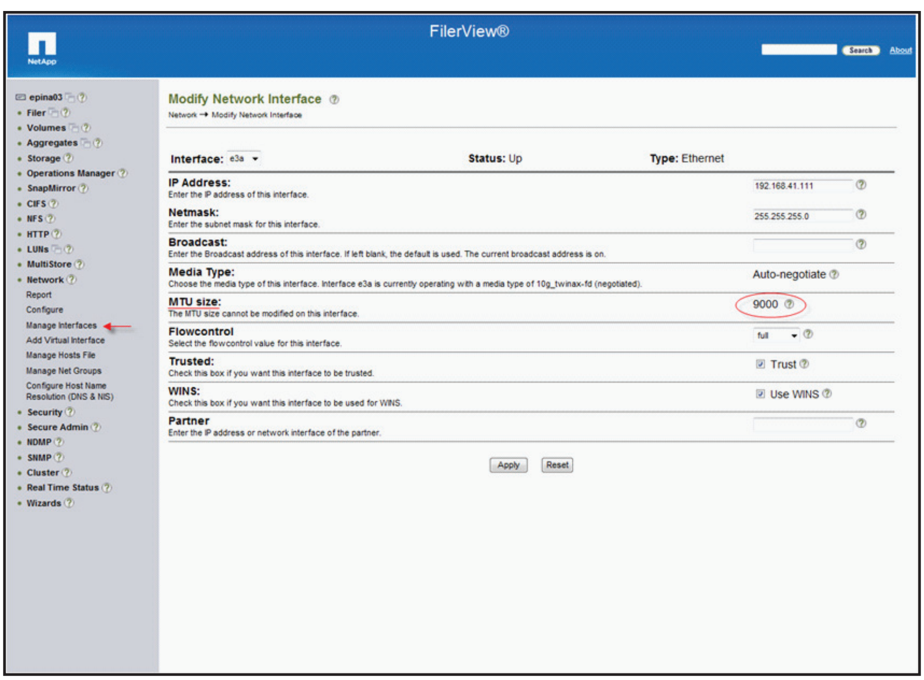


Figure 111

3. To view all network interface parameter settings, select "Show All Interface Details" from the Manage Network Interface window.

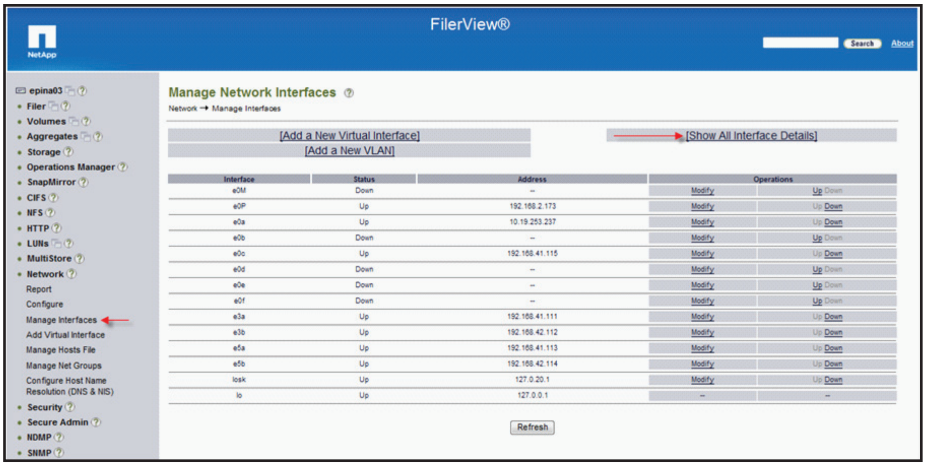


Figure 112

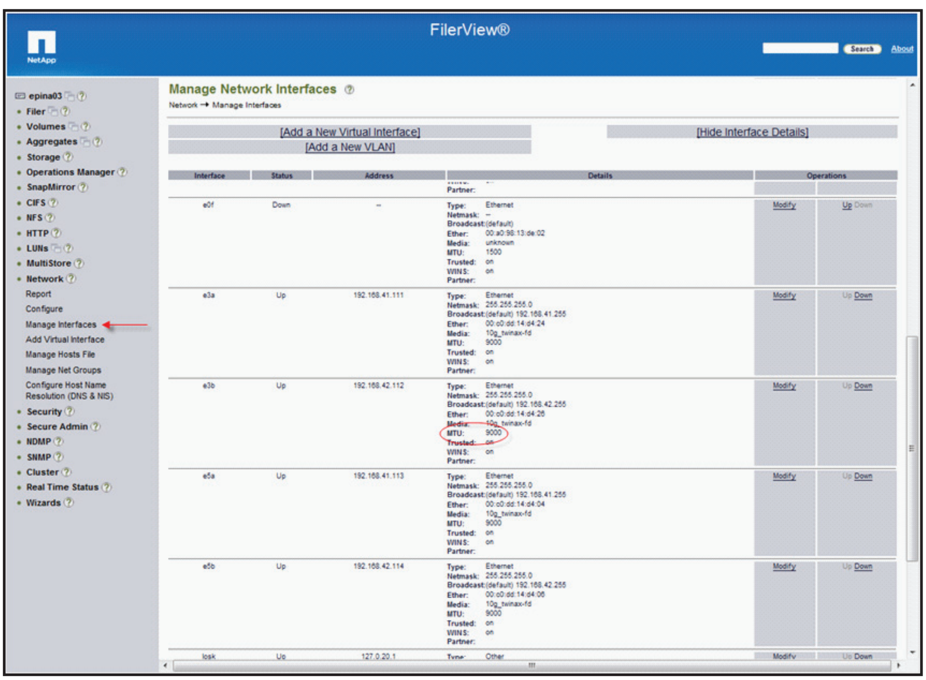


Figure 113

Endnotes

1. Source: Intel estimates as of January 2010. Performance comparison using SPECjbb*2005 bops (business operations per second). Results have been estimated based on internal Intel analysis and are provided for informational purposes only. Any difference in system hardware or software design or configuration may affect actual performance.

2. IDC, "Server Workloads Forecast," 2009.

3. IDC, "The Internet Reaches Late Adolescence," December 2009.

4. 8x Network: 800 terabytes per second of IP traffic estimated on internal Intel analysis "Network Supply/Demand 2010-2020" forecast; 16x Storage: 60 exabytes of data stored from Barclays Capital "Storage Bits" September 2009, extrapolation by Intel for 2015; 20x Compute: Intel internal long-range planning forecast.

5. IDC WW Storage Systems Tracker, December 2009

6. "Multiple Connections per Session (MCS) support is defined in the iSCSI RFC to allow multiple TCP/IP connections from the initiator to the target for the same iSCSI session. This is iSCSI protocol specific. This allows I/O to be sent over either TCP/IP connection to the target. If one connection fails, another connection can continue processing I/O without interrupting the application. Note that not all iSCSI targets support MCS."

To learn more about deployment of cloud solutions,
visit www.intel.com/cloudbuilders

Printed in USA 1010/BA/PRW/PDF Please Recycle 324431-002 US